DNA Damage Recognition Mechanism of Xeroderma Pigmentosum Group C Protein in Human Nucleotide Excision Repair

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Olivier Maillard

von

Cressier FR

Promotionskomitee

Prof. Dr. Josef Jiricny (Vorsitz)

Prof. Dr. Hanspeter Naegeli (Leitung der Dissertation)

Prof. Dr. Michael Hottiger

Zürich, 2007

Dedicated to my Family

Table of Contents **Page**

Chapter I

Introduction and Aim of the Thesis

Xeroderma pigmentosum complementation group C (XPC) protein is the key DNA damage recognition factor that initiates the global genome pathway of mammalian nucleotide excision repair. To serve this unique function, XPC must be able to recognize a broad spectrum of bulky DNA lesions, including those induced by UV irradiation [cyclobutane pyrimidine dimers and (6-4) photoproducts] and many different carcinogen-DNA adducts. However, the molecular mechanism by which this versatile sensor of DNA damage detects a wide range of bulky base lesions is not understood (Friedberg et al., 2006).

The human *XPC* gene product encodes a hydrophilic polypeptide of 940 amino acids, which migrates with an electrophoretic mobility of about 125 kDa in denaturing gels (Masutani et al., 1994). In soluble cell lysates, the XPC protein is found in complexes with HR23B, a 58-kDa homolog of the yeast nucleotide excision repair protein RAD23 (Masutani et al, 1994), and centrin-2 (CTN2), a 18-kDa calcium-binding protein (Araki et al, 2001). XPC protein on its own possesses DNA-binding activity (Uchida et al., 2002), whereas the HR23B and centrin-2 partners exert accessory functions in stabilizing the complex and stimulating its repair activity (Masutani et al., 1997; Ng et al., 2003).

XPC protein alone or in conjunction with HR23B has been shown to bind preferentially to damaged DNA substrates containing, for example, the UV light-induced (6-4) photoproducts, DNA adducts induced by the alkylating agent acetylaminofluorene (AAF) or cisplatin intrastrand crosslinks (Sugasawa et al, 1998; Batty et al, 2000; Kusumoto et al, 2001). Since there is no common chemical or structural motif among the different DNA adducts recognized by XPC protein, the recognition mechanism is not explainable by a classical "lock and key" recognition scheme (Sancar, 1996; Wood, 1997; de Laat et al., 1999). These early observations prompted the hypothesis that XPC protein may detect a particular conformational distortion imposed on the DNA double helix by the different adducts that are processed by the nucleotide excision repair system. Thus, the structural determinants for the recruitment of XPC protein to such DNA lesions have been further probed with artificial DNA substrates, thus revealing an affinity of the XPC subunit for DNA sites that deviate from the canonical Watson-Crick geometry, including a 6-nucleotide mispaired region, mismatched bubbles involving 3 to 5 nucleotides, or single-stranded DNA (Sugasawa et al., 2001; Hey et al., 2002; Sugasawa et al., 2002).

In 2002, Uchida et al. reported that they have been able to narrow down the DNA-binding domain of XPC to a region of 137 amino acids (codons 607-742) within its evolutionary conserved carboxy-terminal half. Because most mutated *XPC* alleles in xeroderma pigmentosum families lead to premature terminations as a result of frameshifts, nonsense mutations, deletions, insertions or aberrant splicing, only one single substitution, which causes a Trp690Ser change, has been identified in this evolutionary conserved region of XPC protein implicated in the interaction with DNA (Chavanne et al., 2000). Although the loss of this aromatic side chain maps to the presumed DNA-binding domain, its consequence with respect to substrate recognition in the nucleotide excision repair pathway is unknown. As a consequence, I performed a mutational screen to analyze the general role of conserved XPC residues in the detection of DNA lesions.

There is ambiguity over the precise amino acid region of XPC protein involved in the complex formation with HR23B. A two-hybrid study reported by Li and collaborators (1997) mapped the HR23B-interacting region of XPC to residues 776 through 801. In contrast, Uchida et al. (2002) performed a bidirectional truncation study to map the minimal HR23B-interacting region of XPC between amino acids 496 and 734. The N-terminal domain of XPC is also responsible for an interaction with XPA, which appears to be important for the transition from an initial recognition intermediate (involving XPC and TFIIH) to the formation of an ultimate incision complex that includes XPA, RPA and the two endonucleases XPF and XPG. In fact, XPC protein behaves like a “molecular matchmaker” as it initiates the assembly of a repair complex but leaves the DNA substrate before completion of the incision reaction (Wakasugi and Sancar, 1998; Riedl et al., 2003; You et al., 2003). The carboxy-terminal tail of XPC protein mediates the association with TFIIH (residues 816-940) (Uchida et al., 2002).

Additional studies showed that XPC protein is also able to interact with 3-methyladenine DNA glycosylase (Miao et al., 2000) and thymine DNA glycosylase (Shimizu et al., 2003). These unexpected findings suggest that the XPC-HR23B-CTN2 complex may constitute a platform not only for the loading of nucleotide excision repair factors onto damaged DNA, but also for the recruitment of a battery of other DNA repair enzymes such as for example different DNA glycosylases involved in the base excision repair pathway. Taken together, these reports emphasize the importance of understanding which domain of XPC is involved in DNA damage

recognition and of identifying the mechanism by which XPC protein discriminates between the native double helix and damaged DNA.

It has been suggested that the recruitment of XPC protein is triggered by conformational distortions of the DNA substrate (Naegeli, 1995; Sugasawa et al., 2002). However, it was not known how this initial factor distinguishes between normal conformational changes of the DNA double helix, induced for example by nucleosome assembly, transcription or other physiologic processes, and the DNA deformation at damaged sites remained elusive. This lack of mechanistic knowledge reflects also the fact that no structural information was available for XPC or one of its eukaryotic XPC homologs. Thus, the purpose of my study was to identify a nucleic acid interaction motif that is responsible for the unique recognition function of XPC protein. The structural and biochemical features of this interaction motif provided an important framework to understand, for the first time, the generic mechanism that is used by XPC protein to detect damaged sites in the genome. I discovered that human XPC protein avoids direct contacts with the damaged bases and, instead, uses a single-stranded DNA-binding motif to recognize the local single-stranded character of the undamaged complementary strand across the lesion site. It is now necessary to perform X-ray crystal analyses or other biophysical studies to confirm this unexpected mode of DNA damage recognition.

References

Araki M, Masutani C, Takemura M, Uchida A, Sugasawa K, Kondoh J, Ohkuma Y, Hanaoka F (2001) Centrosome protein centrin 2/caltractin 1 is part of the xeroderma pigmentosum group C complex that initiates global genome nucleotide excision repair. *J Biol Chem* **276**: 18665-18672

Batty D, Rapic'-Otrin V, Levine AS, Wood RD (2000) Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged sites. *J Mol Biol* **300**: 275-290

Chavanne F, Broughton BC, Pietra D, Nardo T, Browitt A, Lehmann AR, Stefanini M (2000) Mutations in the XPC gene in families with xeroderma pigmentosum and consequences at the cell, protein, and transcript level. *Cancer Res* **60**: 1974-1982.

de Laat WL, Jaspers NG, Hoeijmakers JH (1999) Molecular mechanism of nucleotide excision repair. *Genes Dev* **13**: 768-785

Friedberg EC, Walker GC, Siede W, Wood RD, Schultz RA & Ellenberger T (2006) DNA Repair and Mutagenesis, 2nd edition. ASM Press, Washington DC, USA

Hey T, Lipps G, Sugasawa K, Iwai S, Hanaoka F, Krauss G (2002) The XPC-HR23B complex displays high affinity and specificity for damaged DNA in a true-equilibrium fluorescence assay. *Biochemistry* **41**: 6583-6587

Kusumoto R, Masutani C, Sugasawa K, Iwai S, Araki M, Uchida A, Mizukoshi T, Hanaoka F (2001) Diversity of the damage recognition step in the global genomic nucleotide excision repair in vitro. *Mutat Res* **485**: 219-227

Li L, Lu X, Peterson C, Legerski R (1997) XPC interacts with both HHR23B and HHR23A in vivo. *Mutat Res* **383**: 197-203

Masutani C, Araki M, Sugasawa K, van der Spek PJ, Yamada A, Uchida A, Maekawa T, Bootsma D, Hoeijmakers JH, Hanaoka F (1997) Identification and characterization of XPC-binding domain of hHR23B. *Mol Cell Biol* **17**:6915-6923

Masutani C, Sugasawa K, Yanagisawa J, Sonoyama T, Ui M, Enomoto T, Takio K, Tanaka K, van der Spek PJ, Bootsma D, Hoeijmakers JH, Hanaoka F (1994) Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23. *EMBO J* **13**: 1831-1843

Miao F, Bouziane M, Dammann R, Masutani C, Hanaoka F, Pfeifer G, O'Connor TR (2000) 3-Methyladenine-DNA glycosylase (MPG protein) interacts with human RAD23 proteins. *J Biol Chem* **275**: 28433-28438

Naegeli H (1995) Mechanisms of DNA damage recognition in mammalian nucleotide excision repair. *FASEB J* **9**: 1043-1050

Ng JM, Vermeulen W, van der Horst GT, Bergink S, Sugasawa K, Vrieling H, Hoeijmakers JH (2003) A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. *Genes Dev* **17**: 1630-1645

Riedl T, Hanaoka F, Egly JM (2003) The comings and goings of nucleotide excision repair factors on damaged DNA. *EMBO J* **22**: 5293-5303

Sancar A (1996) DNA excision repair. *Annu Rev Biochem* **65**: 43-81

Shimizu Y, Iwai S, Hanaoka F, Sugasawa K (2003) Xeroderma pigmentosum group C protein interacts physically and functionally with thymine DNA glycosylase. *EMBO J* **22**: 164-173

Sugasawa K, Shimizu Y, Iwai S, Hanaoka F (2002) A molecular mechanism for DNA damage recognition by the xeroderma pigmentosum group C protein complex. *DNA Repair* **1**: 95-107

Sugasawa K, Okamoto T, Shimizu Y, Masutani C, Iwai S, Hanaoka F (2001) A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes Dev* **15**: 507-521

Sugasawa K, Ng JM, Masutani C, Iwai S, van der Spek PJ, Eker AP, Hanaoka F, Bootsma D, Hoeijmakers JH (1998) Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Mol Cell* **2**: 223-232

Uchida A, Sugasawa K, Masutani C, Dohmae N, Araki M, Yokoi M, Ohkuma Y, Hanaoka F (2002) The carboxy-terminal domain of the XPC protein plays a crucial role in nucleotide excision repair through interactions with transcription factor IIH. *DNA Repair* **1**: 449-461

Wakasugi M, Sancar A (1998) Assembly, subunit composition, and footprint of human DNA repair excision nuclease. *Proc Natl Acad Sci USA* **95**: 6669-6674

Wood RD (1997) Nucleotide excision repair in mammalian cells. *J Biol Chem* **272**: 23465–23468

You JS, Wang M, Lee SH (2003) Biochemical analysis of the damage recognition process in nucleotide excision repair. *J Biol Chem* **278**: 7476-7485

Chapter II

An Aromatic Sensor with Aversion to Damaged Strands confers Versatility to DNA Repair

An Aromatic Sensor with Aversion to Damaged Strands Confers Versatility to DNA Repair

Olivier Maillard, Szilvia Solyom, Hanspeter Naegeli*

Institute of Pharmacology and Toxicology, University of Zurich–Vetsuisse, Zurich, Switzerland

It was not known how xeroderma pigmentosum group C (XPC) protein, the primary initiator of global nucleotide excision repair, achieves its outstanding substrate versatility. Here, we analyzed the molecular pathology of a unique Trp690Ser substitution, which is the only reported missense mutation in xeroderma patients mapping to the evolutionary conserved region of XPC protein. The function of this critical residue and neighboring conserved aromatics was tested by site-directed mutagenesis followed by screening for excision activity and DNA binding. This comparison demonstrated that Trp690 and Phe733 drive the preferential recruitment of XPC protein to repair substrates by mediating an exquisite affinity for single-stranded sites. Such a dual deployment of aromatic side chains is the distinctive feature of functional oligonucleotide/oligosaccharide-binding folds and, indeed, sequence homologies with replication protein A and breast cancer susceptibility 2 protein indicate that XPC displays a monomeric variant of this recurrent interaction motif. An aversion to associate with damaged oligonucleotides implies that XPC protein avoids direct contacts with base adducts. These results reveal for the first time, to our knowledge, an entirely inverted mechanism of substrate recognition that relies on the detection of single-stranded configurations in the undamaged complementary sequence of the double helix.

Citation: Maillard O, Solyom S, Naegeli H (2007) An aromatic sensor with aversion to damaged strands confers versatility to DNA repair. PLoS Biol 5(4): e79. doi:10.1371/journal.pbio.0050079

Introduction

One of the most formidable challenges in DNA metabolism is that faced by the initiator of the nucleotide excision repair reaction as it locates damaged sites in the context of a large excess of mostly undamaged residues. This challenge is further complicated by an astounding diversity of target lesions, including cyclobutane pyrimidine dimers and pyrimidine–pyrimidone (6–4) photoproducts induced by UV (ultraviolet) light, bulky DNA adducts generated by electrophilic chemicals [1–4], a subset of oxidative products [5–7], and certain protein-DNA crosslinks [8]. Molecular defects in this versatile nucleotide excision repair response cause autosomal recessive disorders in humans such as xeroderma pigmentosum (XP) or Cockayne syndrome [9–11]. The XP syndrome, in particular, is characterized by photosensitivity and an extreme predisposition to sunlight-induced skin cancer [12]. In addition to cutaneous abnormalities, some XP patients also develop internal tumors [13] or neurologic complications leading to DeSanctis–Cacchione syndrome [14]. Individuals affected by XP are classified into seven repair-deficient complementation groups designated XP–A through XP–G [15].

The nucleotide excision repair response is separated in two pathways. Global genome repair (GGR) activity is responsible for the excision of DNA lesions across all nucleotide sequences, whereas transcription-coupled repair removes offending lesions only from the transcribed strand of active genes [16,17]. A principal difference between these pathways resides in the initial detection of DNA damage. During transcription-coupled repair, elongation of the RNA polymerase II complex is blocked by abnormal residues, thereby inducing the assembly of repair complexes [18]. In contrast, the GGR machinery is dependent on the initial recognition of damaged sites by XPC protein, which constitutes a universal sensor of bulky lesions [19,20]. Recent studies showed that XPC is also required for histone modifications in response to bulky lesion formation, presumably to facilitate chromatin remodeling [21,22]. It has been suggested that the recruitment of XPC protein is triggered by distortions of the DNA substrate [23–25], but how this initial factor distinguishes between normal conformations of the double helix, induced by nucleosome assembly, transcription or other physiologic processes, and the DNA deformation at damaged sites remained elusive. This lack of mechanistic knowledge reflects the fact that no structure is available for any XPC homolog. Thus, the purpose of this study was to identify a nucleic acid interaction motif that is responsible for the unique recognition function of XPC protein.

The human *XPC* gene encodes a polypeptide of 940 amino acids that exists as a complex with centrin 2, a centrosomal protein, and HR23B, one of two mammalian homologs of yeast RAD23. XPC protein itself possesses DNA-binding activity, whereas the centrin 2 and HR23B partners exert accessory functions [26,27]. Uchida et al. [28] have been able to narrow down the DNA-binding domain of XPC to a region of 137 amino acids (codons 607–742) within its evolutionary

Academic Editor: John Tainer, Scripps Research Institute, United States of America

Received July 11, 2006; **Accepted** January 16, 2007; **Published** March 13, 2007

Abbreviations: GFP, green-fluorescent protein; GGR, global genome repair; MBP, maltose-binding protein; OB-fold, oligonucleotide/oligosaccharide-binding fold; RPA, replication protein A; SD, standard deviation; *Sf9, Spodoptera frugiperda*; UV, ultraviolet; XP, xeroderma pigmentosum; XPC, xeroderma pigmentosum group C

* To whom correspondence should be addressed. E-mail: naegelih@vetpharm.unizh.ch

Author Summary

DNA is constantly exposed to damaging agents such as ultraviolet light, carcinogens, or reactive metabolic byproducts causing thousands of DNA lesions in a typical human cell every hour. To prevent irreversible mutations, many of these different lesions are eliminated by a DNA repair system known as "nucleotide excision repair." Repair is initiated by the XPC protein, which recognizes damaged sites in the DNA double helix. Here, we describe how the XPC protein probes the way in which the two DNA strands are aligned, and how a recurrent protein motif, termed oligonucleotide/oligosaccharide-binding fold, is used to detect dynamic fluctuations of DNA in the lesion containing regions. We show that XPC interacts preferentially with the undamaged strand opposite the lesion sites and conclude that XPC protein adopts an entirely indirect recognition mechanism to be able to detect a nearly infinite spectrum of DNA lesions.

conserved carboxy-terminal half. Because most mutated *XPC* alleles in xeroderma pigmentosum families lead to premature terminations as a result of frameshifts, nonsense mutations, deletions, insertions or aberrant splicing, only one single substitution, which causes a Trp690Ser change, has been identified in the evolutionary conserved region of XPC protein [29]. Although the loss of this aromatic side chain maps to the presumed DNA-binding domain, its consequence with respect to substrate recognition in the GGR pathway is unknown, prompting a mutational screen to analyze the general role of conserved XPC residues in the detection of DNA lesions. This study disclosed an aromatic hot spot, consisting of Trp690 and Phe733, which mediates an affinity for the single-stranded character of target sites but with an astonishing aversion to associate with damaged DNA strands. A dual system of aromatics that stack with individual unpaired bases of single-stranded DNA has already been identified in RPA (replication protein A), breast cancer susceptibility 2 protein, and many other single-stranded DNA-binding factors [30–33]. Therefore, our results point to a counterintuitive mechanism of damage recognition by which XPC protein avoids direct contacts with bulky lesions but, instead, probes the local susceptibility of intact nucleotides, on the opposite side of the double helix, to adopt a single-stranded configuration. The spontaneous Trp690Ser point mutation associated with the XP syndrome interferes with this inverted mode of substrate discrimination.

Results

Identification of Evolutionary Conserved Aromatic Residues

The human XPC sequence has been aligned [34] with its homologs from mouse, rat, *Drosophila melanogaster*, *Trypanosoma cruzi*, yeast, and *Arabidopsis thaliana* to identify potential consensus motifs in a region that includes the presumed DNA-binding domain [28]. This sequence alignment demonstrates that Trp690, mutated in an XP family, is maintained from lower eukaryotes to plants and mammals. The only exception is provided by one of the two homologs in *Schizosaccharomyces pombe*, where the regular Trp at this position is replaced by another aromatic residue (Figure 1). The molecular function of an obligatory aromatic side chain at codon 690 was tested by a systematic comparison with all other evolutionary conserved aromatics that were identified in the same portion of human XPC protein, i.e., between codons 531 and 742. Also, the effects of these mutations were evaluated in relation to the substitution of other conserved residues with varying side chains. Figure 1 shows the positions in the presumed DNA-binding domain that have been selected for site-directed mutagenesis and highlights their degree of conservation among eukaryotes.

Conserved Aromatics Are Critical Determinants of XPC Function

A host cell reactivation assay was used to monitor the DNA repair proficiency of XPC mutants in human cells [35]. XP–C fibroblasts, which fail to express XPC protein, were transiently transfected with a dual luciferase reporter system accompanied by an expression vector coding for human XPC protein or the different mutants. The reporter construct, which carries a firefly luciferase gene, was damaged by exposure to UV light (254 nm; 1000 J/m^2) and supplemented with an unirradiated control vector that expresses the *Renilla* luciferase. Following varying repair times, firefly luciferase activity was determined in cell lysates and normalized against the internal *Renilla* standard.

Due to the repair defect of XP–C cells, transcription of the reporter gene was suppressed by persistent UV lesions, resulting in reduced firefly luciferase activity. However, DNA repair and, hence, firefly luciferase expression was restored following transfection with pcXPC, demonstrating that the genetic defect of XP–C fibroblasts is corrected by wild-type XPC protein (Figure 2A). In contrast, expression of the reporter gene was not rescued when the same XP–C cells were transfected with the empty vector pcDNA (Figure 2B). The residual background activity (~15% of wild-type control), observed in the presence of these empty vectors, is likely due to the transcription-coupled repair process, which operates independently of XPC. In part, this residual activity may also result from a minor fraction of plasmids remaining free of bulky UV lesions in the luciferase reporter sequence.

The firefly luciferase production was not restored when, instead of XPC, XPA protein was expressed in XP–C fibroblasts (Figure 2B), thus demonstrating the specificity of our host cell reactivation system. Also, the firefly luciferase production was inhibited when the XPC sequence was modified to carry the Trp690Ser mutation responsible for clinical manifestations of the XP syndrome (Figure 2B). Nearly identical results were obtained by transfecting the cells with vector pXPC–GFP, which drives the expression of wild type or mutated XPC sequences fused, on their carboxy-terminal side, to green fluorescent protein (GFP). As expected, no complementation of the repair defect was detected upon expression of GFP alone using the corresponding control vector (Figure 2B).

The relative luciferase activity indicative of DNA repair was determined in the presence of each site-directed mutant, and the results were reported as the percentage of wild-type complementation after deduction of background luciferase expression. Initially, the aromatic side chains of conserved Phe, Trp, and Tyr residues were eliminated by Ala substitutions (Figure 2C). In most cases, the excision-repair proficiency of XPC protein was only marginally diminished by these Phe→Ala, Trp→Ala, or Tyr→Ala changes. However, point mutations at the conserved codons 531, 542, 585, 690,

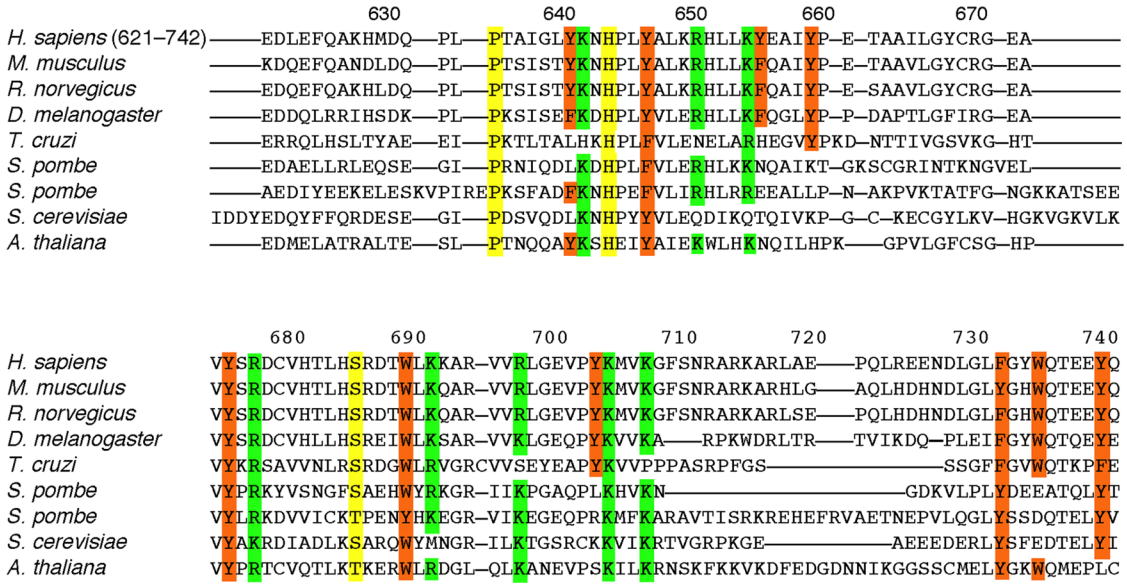


Figure 1. Evolutionary Conserved Residues in the Proposed DNA-Binding Domain of XPC Protein

Sequence comparison between eukaryotic XPC homologs. There are two homologous genes in *S. pombe*. Amino acids targeted by site-directed mutagenesis are highlighted. Y, F, W, aromatic (orange); K, L, positively charged (green); P, H, S, other highly conserved positions (yellow).
doi:10.1371/journal.pbio.0050079.g001

and 733 resulted in a substantial (>50%) reduction of excision activity, and the residual DNA repair observed with these mutants is similar to the low level of complementation promoted by the Trp690Ser allele (Figure 2C). All these mutants displayed essentially the same repair deficiency when reexamined as GFP fusion products (unpublished data).

The more sensitive codons 531, 542, 585, 690, and 733 were further tested by converting the respective aromatics to different amino acids with varying properties. In all cases, the luciferase activity reflecting DNA excision repair was strongly reduced regardless of whether the aromatics were replaced by the aliphatic side chain of Ala, the hydrogen moiety of Gly, or the hydrophilic side chain of Ser (Figure 2D). These results imply that the loss of activity conferred by these XPC mutations is primarily a consequence of the missing aromatic residue rather than being dependent on the properties of the newly introduced substituent.

Basic amino acids frequently make contacts with the phosphate moieties of the DNA backbone. Thus, evolutionary conserved Lys and Arg residues, located between codons 595 and 708 of the human XPC protein, were targeted by site-directed mutagenesis. The positively charged side chains were eliminated by changing the respective residues to Gly, but none of the resulting Lys→Gly or Arg→Gly substitutions were able to perturb the XPC function (Figure 2E). In addition, absolutely conserved amino acids in the center of the putative DNA-binding domain of human XPC protein were changed to Ala residues. The resulting Pro635Ala, His644Ala, and Ser686Ala substitutions reduced the luciferase activity to a moderate degree but, interestingly, none of these mutants reached the low residual repair level observed after removal of an aromatic side chain at position 690 or 733 (Figure 2F).

Normal Expression and Cellular Localization of Repair-Deficient XPC Mutants

The cellular XPC content was monitored by immunoblot analysis of XP–C fibroblasts harvested 15 h after transient transfections with vector pcXPC, promoting the expression of human XPC alone, or vector pXPC–EGFP translating to the production of XPC as a GFP fusion protein. In both cases, a quantitative comparison of protein levels demonstrated that the Trp690Ser and Trp690Gly mutants were expressed in human fibroblasts to similar levels as the wild-type counterpart (Figure 3A and 3B). Moreover, the repair-deficient mutants with Ala substitutions at codons 531, 542, 585, 690, and 733 were detected in human fibroblasts in nearly identical amounts as wild-type XPC protein (Figure 3C). Thus, the repair deficiency observed by substituting these conserved aromatics is not a consequence of reduced XPC expression or enhanced degradation.

The GFP fusion partner was exploited to perform fluorescence microscopy studies. A time course experiment with the wild-type sequence demonstrated that expression of the XPC–GFP fusion increases during incubation periods of 18 h after transfection, with a cellular localization that is predominantly restricted to the nucleus (Figure 3D). Control cells transfected with vector pGFP demonstrated that GFP alone displays a more diffuse distribution extending to both the cytoplasma and nucleus (Figure 3E). However, the strong nuclear localization is reestablished after expression of GFP fused to the Trp690Ser mutant (Figure 3F). A similar level of fluorescence with the same characteristic nuclear localization was recorded for each of the repair-defective Ala mutants (Figure 3G). These results demonstrate that the repair deficiency of these tested mutants is not due to defective translocation into the nuclear compartment.

XPC Protein Displays a Single-Stranded DNA-Binding Motif

The wild-type XPC polypeptide was coupled to maltose-binding protein (MBP), produced in *Spodoptera frugiperda* (*Sf*9) cells and purified to homogeneity by nickel and heparin affinity chromatography. MBP was chosen as a fusion partner to promote solubility and proper folding [36]. Another

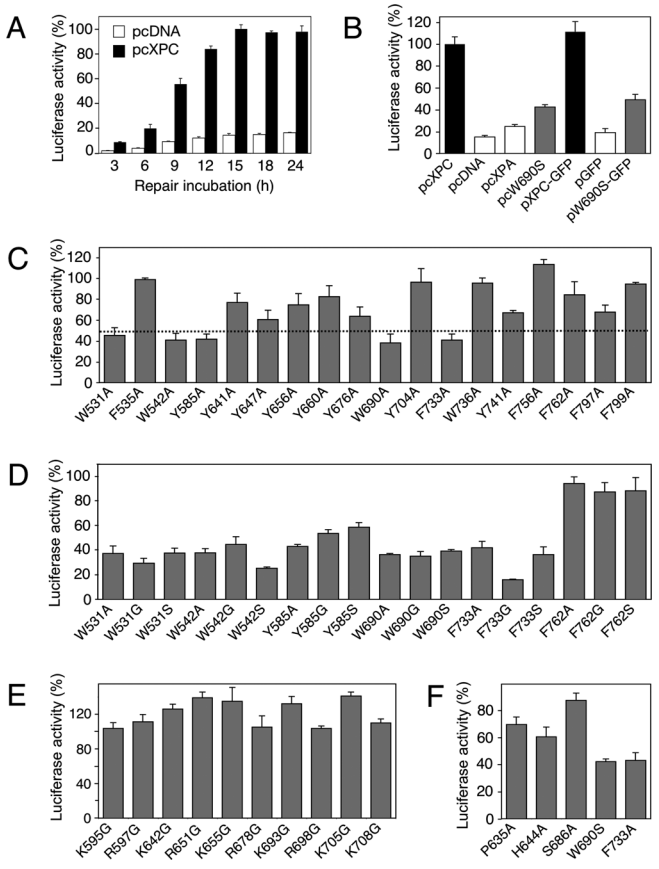


Figure 2. Screening for Repair-Deficient XPC Mutants

(A) Time course of host cell reactivation assay. Human XP–C fibroblasts were transfected with an expression vector coding for wild-type XPC (pcXPC) or the empty control vector (pcDNA). Repair complementation was assessed after the indicated times by monitoring luciferase expression from a UV-irradiated reporter construct. Excision is reported as the percentage of wild-type activity after 15 h of incubation (± SD).

(B) Specificity of the repair assay. XP–C fibroblasts were transfected with expression vectors coding for wild-type XPC (pcXPC or pXPC–GFP), wild-type XPA (pcXPA), the control vectors (pcDNA and pGFP), or vectors containing the Trp690Ser mutant sequences (pcW690S or pW690S–GFP). Excision is reported as the percentage of wild-type activity (15-h incubations).

(C) Deletion of aromatic side chains. XP–C fibroblasts were transfected with vector pcXPC carrying the indicated mutations. DNA repair is expressed as the percentage of wild-type complementation (15-h incubations) after deduction of background luciferase activity obtained with the control vector. The dashed line indicates a threshold of 50% reduction in repair activity.

(D) Replacement of aromatic residues by amino acids of different properties.

(E) Deletion of the positively charged side chains of conserved Lys or Arg residues.

(F) Ala substitutions of absolutely conserved positions in the center of the DNA-binding domain.

doi:10.1371/journal.pbio.0050079.g002

advantage of the MBP tag is that, on its own, it lacks DNA-binding activity [37]. On sodium dodecylsulfate gels, the final fraction of the MBP–XPC fusion product migrated as a single band with an apparent molecular weight of ~170 kDa, which corresponds to the expected size of the 125-kDa XPC protein linked to the 43-kDa MBP moiety (Figure 4A).

Conflicting results regarding the affinity of XPC protein for DNA substrates of different lengths and conformations have emerged. Oligonucleotides with fewer than 60 base pairs resulted in weakened binding and reduced damage selectivity [25,38,39]. As a consequence, we employed radiolabeled duplexes of 65 base pairs to monitor DNA binding in electrophoretic mobility shift assays. The nucleotide sequence was designed to contain neighboring pyrimidines for the formation of UV-induced dimers. Thus, the double-stranded substrates were UV irradiated (254-nm wavelength) to test the DNA damage selectivity of purified XPC fusion products. As expected from previous reports [40,41], an increased affinity of XPC for UV-irradiated duplexes, over the unirradiated control, was detected when the binding reactions were supplemented with an excess of undamaged competitor DNA, i.e., under conditions of limiting protein (Figure 4B).

In addition to this known affinity for UV-irradiated duplexes, we observed that XPC protein exhibits an extraordinary preference for binding to single-stranded 65-mer oligonucleotides over undamaged double-stranded fragments of the same length (Figure 4C). These results obtained with relatively long oligomeric substrates imply that the XPC subunit fits the classic definition of a single-stranded DNA-binding protein. Because shorter duplexes are more prone to spontaneous denaturation, generating regions of single-stranded DNA, the preference of XPC protein for binding to single strands over double-stranded DNA is abrogated by reducing the oligonucleotide length to 40 residues or fewer (unpublished data). This effect of substrate length provides a possible explanation for the diverging results of previous studies where the damage selectivity of XPC protein had not been attributed to an affinity for single-stranded DNA conformations [23,40]. A striking bias for single-stranded DNA is further supported by competition assays showing that the binding of XPC protein to UV-irradiated 65-mer duplexes is sensitive to the addition of 65-mer single strands (Figure 4D). Conversely, when the competitor consisted of double-stranded plasmids, an excess of heavily UV-irradiated DNA was necessary to reduce the binding of XPC protein to single-stranded oligonucleotides (Figure 4E).

Subsequently, we observed that the high-affinity association of XPC protein with DNA single strands was progressively reduced upon UV irradiation of the oligonucleotide substrate (Figure 4F). Interestingly, the UV dose of 600 J/m^2 is expected to yield a damage frequency of <1 photoproduct/oligonucleotide molecule (40), yet this low level of radiation was sufficient to reduce the single-stranded DNA-binding activity of XPC protein by ~50%. Higher UV doses further suppressed the single-stranded DNA-binding activity to marginal levels (Figure 4G), indicating that bulky lesions collide with the ability of XPC protein to form complexes with DNA oligonucleotides. Taken together, we conclude that XPC protein is recruited to target sites by virtue of its characteristic preference for deoxyribonucleotide sequences that adopt a single-stranded conformation. Surprisingly, this sensor protein associates preferentially with undamaged strands but rejects direct interactions with damaged strands.

The Trp690Ser Substitution Confers Defective DNA Binding

Two different strategies were used to test the ability of XPC mutants to interact with single-stranded DNA substrates. First, MBP–XPC fusion products were expressed in *Sf*9 cells, and the respective cell lysates were incubated with single-stranded DNA immobilized on agarose beads. After 2 h-

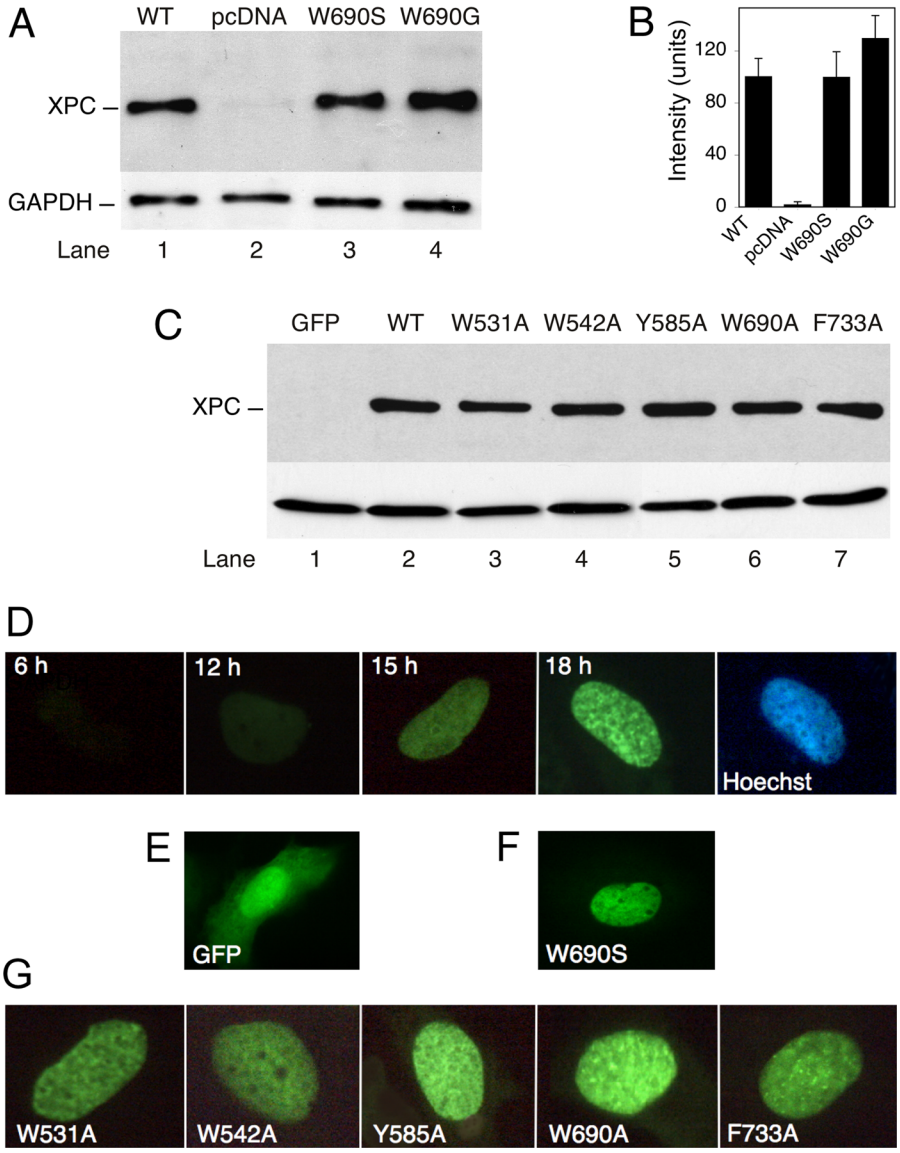


Figure 3. Normal Cellular Expression and Localization of Repair-Deficient XPC Mutants

(A) Immunoblot analysis of XP–C fibroblasts transfected with pcXPC vectors coding for wild-type protein or repair-deficient mutants. Soluble cell lysates (20 μg) were separated on polyacrylamide gels and probed for XPC protein using a specific monoclonal antibody directed against the C-terminal sequence of human XPC. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was recorded as the internal standard. Lane 2: XP–C fibroblasts transfected with the pcDNA control vector.

(B) Densitometric quantification of three to five independent experiments. The intensity of immunoreactive bands corresponding to XPC protein was normalized against the glyceraldehyde-3-phosphate dehydrogenase standard and reported as the percentage of the wild-type signal (± SD).

(C) Immunoblot analysis of XP–C fibroblasts transfected with pGFP (lane 1), or the vectors coding for wild-type (lane 2) or mutant XPC proteins (lanes 3–7) fused to GFP. The primary antibody was directed against GFP.

(D) Time course of fluorescent fusion protein expression in XP–C fibroblasts transfected with pXPC–GFP containing the wild-type sequence. Nuclei were stained with Hoechst reagent.

(E) Distribution of GFP in XP–C fibroblasts.

(F and G) Representative images demonstrating the nuclear localization of repair-deficient mutants 15 h after transfection.

doi:10.1371/journal.pbio.0050079.g003

incubations at 4 °C, the fraction of XPC protein in the pellet (bound to DNA) was separated by repeated washing from the free XPC molecules remaining in the supernatant. The extensively washed pellets and the accompanying supernatants were analyzed separately by gel electrophoresis and immunoblotting. Side-by-side comparisons showed that, in the case of the wild-type control, a major proportion (>70%) of XPC protein was recovered in the DNA-agarose pellet when the binding reactions were performed in buffer containing NaCl concentrations of 0.1–0.3 M (Figure 5A, lanes 1–6). If the NaCl concentration was raised to 0.4 M, only ~50% of wild-type protein remained bound to DNA (Figure 5A, lanes 7 and 8). When the ionic strength was further increased, the proportion of XPC protein retained in the DNA pellet was diminished, reflecting a gradual reduction of nucleic acid binding. In the case of the Trp690Ser mutant, the fraction of protein recovered in association with the DNA beads was markedly reduced already in buffer containing 0.1

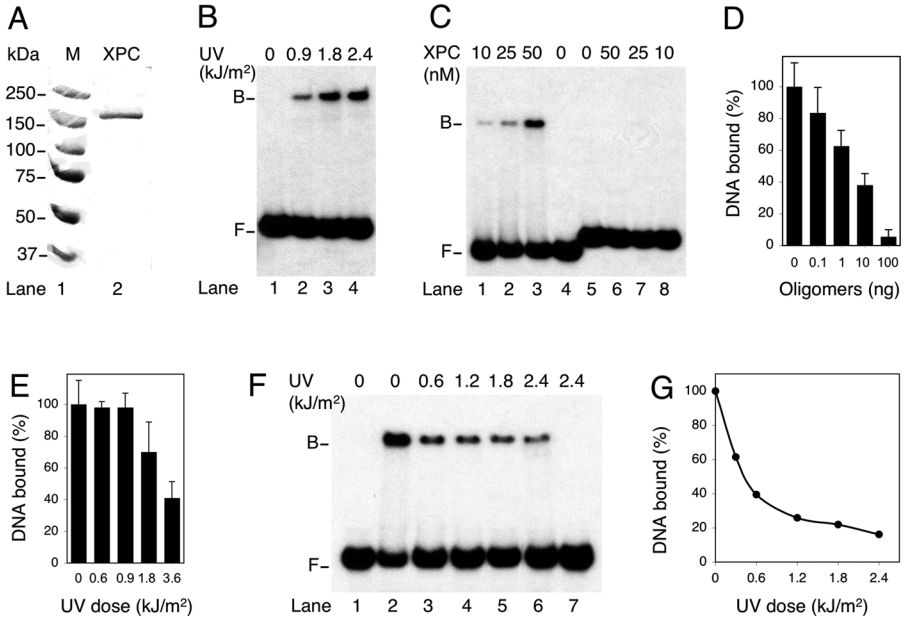


Figure 4. Affinity of XPC Protein for Native Single-Stranded Oligonucleotides

(A) Analysis of MBP–XPC fusion protein by Coomassie staining of a denaturing 8% polyacrylamide gel. Lane 1, markers; lane 2, purified fraction.
(B) Electrophoretic mobility shift assay demonstrating the preference of wild-type XPC protein for UV-irradiated duplexes over the unirradiated control (lane 1). Radiolabeled double-stranded DNA fragments of 65 base pairs (2 nM) were incubated for 30 min with the MBP–XPC fusion product (50 nM) and duplex poly[dI-dC] (10 ng/μl). F, free DNA; B, protein-bound DNA.
(C) Preference of XPC protein for binding to single-stranded 65-mer oligonucleotides (lanes 1–4) relative to undamaged 65-mer duplexes (lanes 5–8).
(D) Competition with single-stranded DNA. Radiolabeled 65-mer duplexes were UV-irradiated (1.8 kJ/m^2) and incubated at a concentration of 2 nM with XPC protein (50 nM), increasing amounts of unlabeled single-stranded oligomers of 65 nucleotides, and duplex poly[dI-dC] (10 ng/μl). The fractions of protein-bound oligomers were determined by electrophoretic mobility shift assay, quantified by laser scanning densitometry, and expressed as the percentage of binding observed in the absence of competitor DNA (± SD).
(E) Competition with double-stranded DNA. Radiolabeled 65-mer oligonucleotides (2 nM) were incubated with XPC protein (50 nM), 100 ng of plasmid DNA (pcDNA) exposed to the indicated UV doses, and duplex poly[dI-dC] (10 ng/μl). The fractions of protein-bound oligomers were quantified and expressed as the percentage of binding determined in the presence of undamaged competitor DNA (± SD).
(F) Suppression of single-stranded DNA binding by UV irradiation. Radiolabeled 65-mer oligonucleotides (2 nM), exposed to the indicated UV doses, were incubated for 30 min with XPC protein (100 nM) and duplex poly[dI-dC] (10 ng/μl). Lane 1: no XPC protein.
(G) Quantification by laser scanning densitometry of two independent experiments performed with UV-irradiated single-stranded oligonucleotides (± SD).
doi:10.1371/journal.pbio.0050079.g004

M NaCl (Figure 5B, lanes 1 and 2). When the NaCl concentration was increased to 0.2 or 0.3 M, the proportion of mutant XPC protein remaining in the DNA pellets was further reduced to ~20% or less (Figure 5B, lanes 3–6). Essentially none of the Trp690Ser mutant remained assembled with DNA when the NaCl concentration was raised to 0.4 M (Figure 5B, lanes 7 and 8). These results show that the Trp690Ser substitution identified in an XP family disrupts the affinity of XPC protein for its DNA substrate.

Trp690 and Phe733 Define an Aromatic Hotspot for Substrate Recognition

All repair-deficient substitutions were expressed as MBP fusion products and tested for their ability to interact with single-stranded DNA immobilized on agarose beads. This systematic comparison was performed in buffer containing 0.3 M NaCl, which corresponds to the ionic strength under which the most pronounced difference was detected between wild-type XPC protein and the Trp690Ser reference. Under these conditions, the three mutants Trp531Ala, Trp542Ala, and Tyr585Ala, which carry Ala substitutions outside the presumed DNA-binding domain, displayed a gradually reduced DNA-binding capacity compared to wild-type XPC protein (Figure 6A), possibly reflecting indirect structural effects on the substrate recognition surface. This gradient of decreasing interactions with DNA culminated in the nearly complete loss of substrate binding in response to the Trp690Ala or Phe733Ala substitution. In both cases, the vast majority of mutant Trp690Ala and Phe733Ala protein appeared as free molecules in the supernatant, and only an insignificant fraction of these two species remained bound to the single-stranded DNA agarose beads (Figure 6A). The Phe762Ala substitution, which yielded only a mild DNA repair defect in the host cell reactivation assay, was included in this nucleic acid-binding screen as an additional control. In full agreement with its in vivo repair proficiency, this Phe762Ala mutant was able to associate with the DNA substrate nearly as efficiently as the wild-type counterpart.

Among the repair-deficient XPC mutants identified in this study, only the Phe733Ala substitution resulted in the same poor DNA-binding activity as the XP mutation at codon 690. Therefore, an independent preparation of this Phe733Ala mutant (Figure 6B, lanes 3 and 4) was reexamined for DNA binding in comparison with newly prepared cell lysates containing the repair-deficient Trp690Ser mutant (lanes 1 and 2), the repair-proficient Phe762Ala derivative, (lanes 5 and 6) as well as the wild-type XPC control (lanes 7 and 8). This control experiment, again carried out in the presence of 0.3 M NaCl, confirmed that the removal of an aromatic side chain at positions 690 and 733 disrupts the DNA-binding

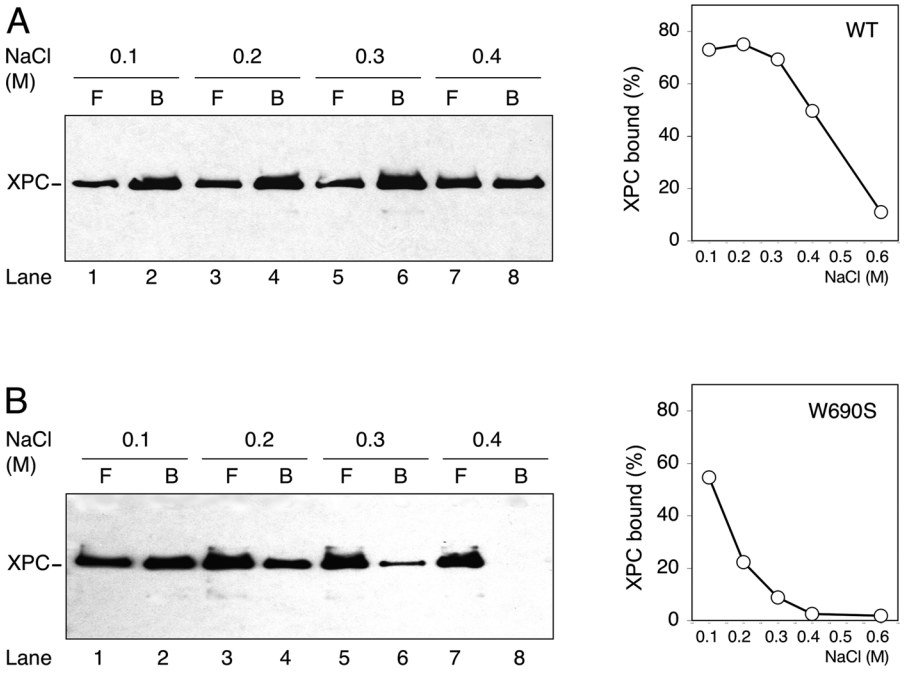


Figure 5. The Trp690Ser Mutant Is Defective in Substrate Binding

(A) Pull-down assays were performed by coincubating *Sf*9 cell lysate (5 μl) containing wild-type XPC protein and 50 μl of single-stranded DNA beads. The binding buffer was supplemented with the indicated concentrations of NaCl. The fractions of free (F) and bound (B) protein were separated and analyzed by gel electrophoresis and immunobloting using specific monoclonal antibodies. The panel on the right provides a quantitative evaluation of three independent binding assays showing the proportion of pulled-down XPC protein at the different ionic strengths.

(B) Pull-down assay with *Sf*9 cell lysate (5 μl) containing the Trp690Ser mutant (left) and quantitative evaluation of three independent experiments (right).

doi:10.1371/journal.pbio.0050079.g005

function of XPC protein. Thus, the molecular defect underlying the prominent repair deficiency of these Trp690 and Phe733 substitutions resides with the inability of the respective mutants to undergo close contacts with the DNA substrate.

Probing of XPC Mutants with Single-Stranded Oligonucleotides

A second experimental strategy, based on defined oligonucleotide probes, was established to confirm that the mutations at codons 690 and 733 confer defective binding to single-stranded DNA. For that purpose, MBP–XPC products were first purified from *Sf*9 cell lysates by immunoprecipitation with anti-MBP antibodies linked to paramagnetic beads. This one-step procedure generated nearly homogenous preparations of MBP–XPC fusion proteins (Figure 7A). Subsequently, the amount of paramagnetic beads was adjusted to include 100 ng of purified protein, translating to a final XPC concentration of 3 nM in each binding reaction. Such purified fractions of wild-type protein or Trp690Ser mutant were incubated with radiolabeled 65-mer single strands and, following 2 h at 4 °C, the oligonucleotides captured by XPC protein were separated from free DNA. After extensive washing, the radioactivity associated with XPC protein on the paramagnetic beads was quantified by scintillation counting. We found substantial binding of wild-type XPC protein to single-stranded oligonucleotides but this interaction was markedly reduced when the Trp690Ser mutant was tested under exactly the same conditions (Figure 7B). Next, the reaction mixtures were adjusted to contain different amounts of protein, thus demonstrating a dose-dependent increase of DNA-binding activity in the presence of wild-type XPC. These dose-dependence experiments confirmed that XPC protein interacts more efficiently with 65-mer heteroduplexes containing a 3-nucleotide bubble than to perfectly homoduplex controls (Figure 7C). The DNA-binding activity was further enhanced by replacing duplex substrates with single-stranded oligonucleotides of the same length (Figure 7C). Finally, these dose-dependent binding assays were used to compare the relative affinity of wild-type and mutant proteins for single-stranded DNA. In contrast to the efficient association of wild-type XPC with 65-mer oligonucleotides, the ability to interact with single-stranded DNA was essentially lost when we tested the mutants carrying an Ala substitution at codon 690 or 733 (Figure 7D). However, in agreement with the different assay of Figure 6, the DNA-binding activity was more moderately affected by a Trp531Ala substitution (Figure 7D). These results support the conclusion that the two aromatic residues Trp690 and Phe733 are critically required for the recognition of single-stranded DNA conformations.

Discussion

The most astounding feature of the GGR machinery is its ability to eliminate a wide diversity of DNA lesions, but how this repair system discriminates anomalous residues against the vast background of normal deoxyribonucleotides is still a focus of intense research, mainly because there is no common chemical motif among the different DNA adducts that would

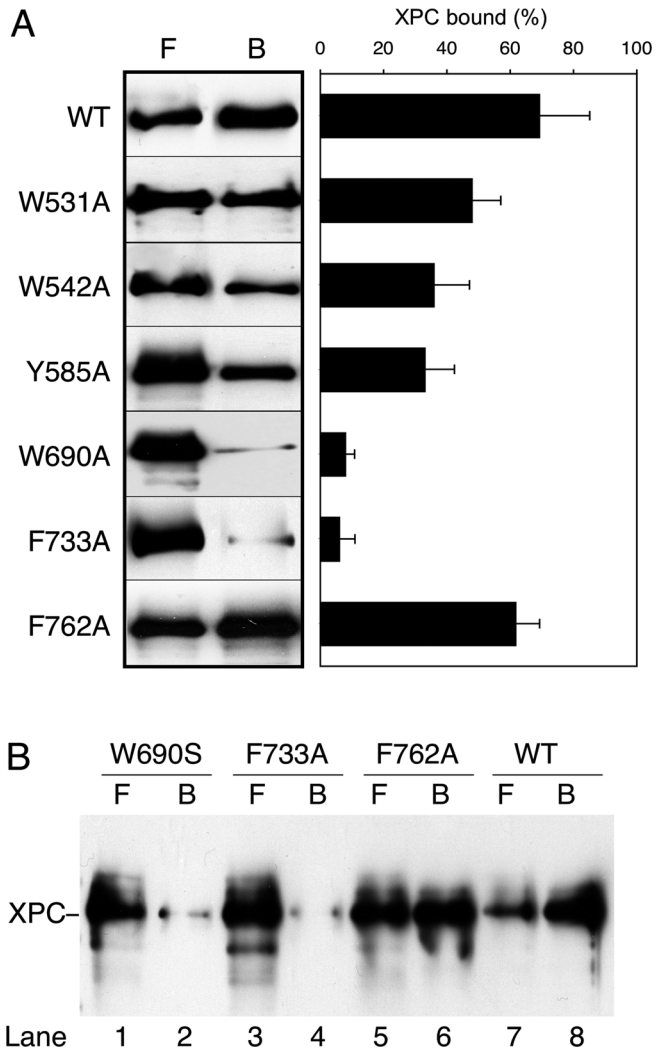


Figure 6. DNA-Binding Deficiency of Trp690 and Phe733 Mutants

(A) Pull-down assays were performed by coincubating *Sf*9 cell lysate (5 μl) containing wild-type XPC or the indicated Ala mutants and single-stranded DNA beads. The binding buffer contained 0.3 M NaCl. The fractions of free (F) and bound (B) protein were separated and analyzed by gel electrophoresis and immunobloting using specific monoclonal antibodies. The panel on the right shows the quantitative evaluation of three independent binding assays (± SD).

(B) Side-by-side comparison of the DNA-binding capacity of wild-type XPC protein (lanes 7 and 8) and the indicated mutants (lanes 1–6).

doi:10.1371/journal.pbio.0050079.g006

account for a classic "lock and key" recognition scheme [1–4]. Our mutagenesis screen designed to probe the mode of action of human XPC protein indicates that this primary initiator of the GGR reaction donates a pair of aromatic side chains (Trp690 and Phe733) to monitor the double helical integrity of DNA and to recognize the local single-stranded character imposed on the undamaged side of the DNA duplex. These novel findings have several important implications with regard to damage recognition and the versatile GGR pathway.

First, the preference of XPC protein for substrates containing a short single-stranded segment, over fully complementary duplexes, provides a truly universal mechanism for the detection of lesion sites. Normally, the native DNA duplex is stabilized by complementary base pairing as well as by stacking interactions between adjacent bases such that, in the absence of damage, the bases are positioned to the interior of the double helix. In contrast, DNA at damaged sites deviates considerably from this canonical Watson–Crick geometry. Bulky adducts often disrupt normal pairing and stacking interactions, thereby lowering the thermal and thermodynamic stability of the duplex, which results in local separation of the complementary strands and exposure of unpaired and unstacked bases on the surface of the double helix, thus generating an abnormal configuration with features that resemble single-stranded DNA [24]. The present equilibrium binding studies as well as kinetic measurements [25], both demonstrating an extraordinary affinity for single-stranded oligonucleotides relative to double-stranded counterparts, imply that only base adducts that destabilize the double helix generate the key molecular signal for recognition by the single-stranded DNA-binding motif of XPC protein.

Second, our results point to an inverted mode of recruitment mediated by an affinity for the undamaged strand of the DNA duplex. In fact, we observed an unfavorable binding of XPC protein to UV-irradiated DNA oligonucleotides compared to undamaged single-stranded counterparts. A similar reduction of oligonucleotide binding has been detected following the introduction of a site-specific cisplatin adduct [25], implying that the interaction of XPC protein with single-stranded DNA is generally disturbed by the presence of adducted, crosslinked, or otherwise aberrant base residues. Thus, the exquisite affinity of XPC protein for single-stranded oligonucleotides, in combination with its aversion to interact with damaged strands, indicates that the recognition step in the GGR pathway is guided by an initial association with the native strand of damaged duplexes (Figure 8A), without ruling out the possibility that XPC protein may ultimately interact with both strands. Such an inverted mode of damage recognition, which is completely independent of the variable chemistry of the lesion sites, accommodates the ability of the GGR machinery to detect a very wide array of DNA adducts. Recently, it has been reported that RPA is equally refractory to interactions with damaged oligonucleotides [42], suggesting a functional analogy between XPC protein and representatives of the large family of single-stranded DNA-binding factors.

Third, the dependence on a dual system of aromatic amino acids indicates a structural basis for the observed similarity between the XPC subunit and known single-stranded DNA-binding proteins. We found that two distinct aromatics in the presumed nucleic acid-binding domain of XPC protein, i.e., Trp690 and Phe733, are more critically involved in the high-affinity interaction with single-stranded configurations than all other conserved residues in the same XPC region. Even mutations affecting the absolutely conserved Pro635, Lys642, His644, Tyr676, Arg678, Ser686, or Lys708, located in the DNA-binding domain, cause less incisive repair deficiencies than the removal of the aromatic side chains at positions 690 and 733. Other aromatic side chains at codons 531, 542, and 585 are similarly required for excision repair activity, but their removal confers more moderate DNA-binding defects. This observation is consistent with a previous report indicating that residues 531–585 are located outside the core DNA-binding domain [28]. The distinctive requirement for a pair of aromatics (Trp690 and Phe733 in the case of XPC) is

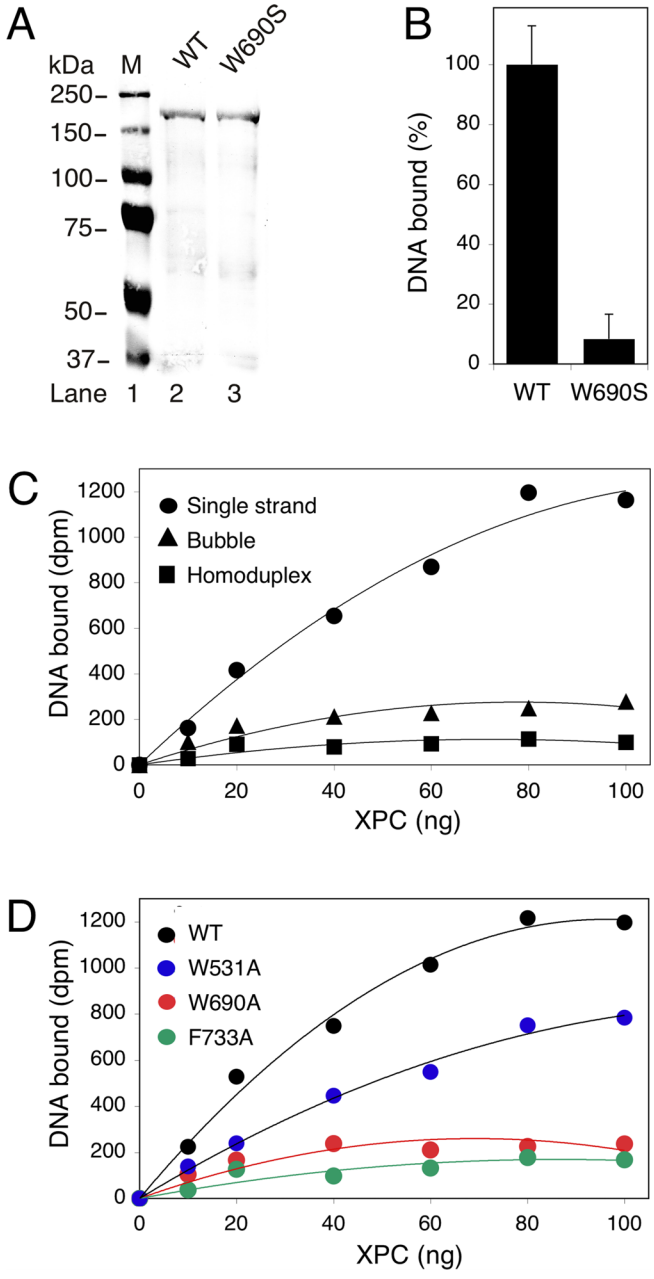


Figure 7. Single-Stranded Oligonucleotide-Binding Defect

(A) Analysis of immunoprecipitated XPC by Coomassie staining of a denaturing 8% polyacrylamide gel. The MBP–XPC fusions were purified from *Sf*9 lysates using monoclonal anti-MBP antibodies linked to paramagnetic beads. Lane 1, markers; lane 2, wild-type MBP–XPC; lane 3, fusion protein containing the Trp690Ser reference mutation.

(B) Binding of wild-type XPC and Trp690Ser mutant to single-stranded oligonucleotides. Immunoprecipitated MBP–XPC protein (100 ng, 3 nM) was incubated with ^{32}P-labeled 65-mer oligonucleotides (2 nM). The DNA molecules captured by XPC protein were separated from the free oligonucleotides and quantified in a scintillation counter. Single-stranded DNA-binding activity (± SD) is reported as the radioactivity immobilized by XPC after deduction of the background binding determined with empty beads.

(C) Differential binding to distinct DNA conformations. Immunoprecipitated MBP–XPC protein (100 ng, 3 nM) was incubated with ^{32}P-labeled substrates (2 nM) consisting of 65-mer homoduplexes, 65-mer heteroduplexes with a central 3-nucleotide bubble, or 65-mer single-stranded oligonucleotides. The DNA molecules captured by XPC protein were separated from free DNA and quantified in a scintillation counter. DNA-binding activity (mean values of two experiments) is reported as the radioactivity immobilized by XPC after deduction of the background binding determined with empty beads.

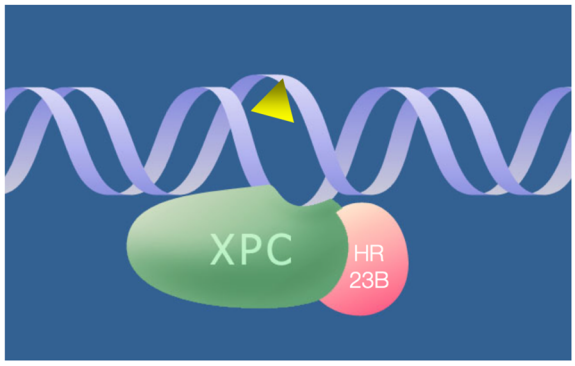
(D) Comparison between wild-type XPC and Ala mutants. Paramagnetic beads containing the indicated amounts of immunopurified MBP–XPC protein were incubated with ^{32}P-labeled 65-mer oligonucleotides (2 nM). DNA associated with XPC protein was separated from the free oligonucleotides and quantified in a scintillation counter. Single-stranded DNA binding activity (mean values of four experiments) is reported as the radioactivity immobilized by XPC after deduction of the background binding to empty beads.

doi:10.1371/journal.pbio.0050079.g007

reminiscent of the OB-fold of many single-stranded DNA-binding proteins [30]. In RPA, for example, four different DNA-binding subdomains with the characteristic OB-fold are responsible for the association with single-stranded substrates [33]. Each of these domains forms a small β-barrel consisting of several short elements of secondary structure connected by loops of variable length [43]. The single-stranded DNA-binding activity of these RPA subdomains correlates with the presence of two structurally conserved aromatics that mediate stacking interactions with closely spaced DNA bases. Other OB-folds in the RPA complex that lack these aromatic side chains fail to contribute to nucleic acid binding [33]. The reiteration of a pattern of two separate aromatics in the DNA-binding domain of XPC protein lends support to the hypothesis that this repair factor may display an analogous structural fold to recognize DNA bases extruded from the double helix, and forced into a single-stranded conformation, as a consequence of bulky lesion formation. The different OB-fold subdomains of RPA range between 110 and 180 amino acids in length. As a minimal DNA-binding fragment of XPC protein has been mapped to a region of 136 amino acids [28], we predict that XPC displays a monomeric variant of this motif to detect the single-stranded character resulting from separation of just one or, depending on the extent of DNA distortion, no more than a few base pairs at lesion sites.

To summarize, XPC protein displays a range of properties that are typical of the OB-fold of single-stranded DNA-binding factors, i.e., an affinity for single-stranded oligonucleotides, an exquisite preference for undamaged strands relative to damaged strands, the pairwise deployment of aromatics for nucleic acid binding, and the ability to interact with single-stranded DNA under conditions of elevated ionic strength. This combination of functional and structural analogies raises the question of whether a common sequence motif may be shared by XPC and known single-stranded DNA-binding proteins. A systematic analysis of the XPC full-length sequence did not reveal any signature that may have predicted its DNA-binding properties [44,45]. However, a homology search focused on the comparison with the growing family of OB-fold proteins showed that the nucleic acid-binding region of XPC protein displays a remarkable similarity to one of the oligonucleotide-binding subdomains of human RPA (Figure 8B). This comparison yielded 27% identity and 73% similarity between the DNA-binding domain of XPC protein and the RPA-B motif situated in the large subunit of the human RPA complex. The sequence homology extends over most of the conserved elements of secondary structure of the RPA-B subdomain and exceeds the 12% identity detected when known OB-folds were aligned according to their high-resolution structure [32]. The same DNA-binding region of XPC also displays a 66% similarity with the OB1 and a 64% similarity with the OB2 motif of breast cancer susceptibility 2 (unpublished data). Thus, the aromatic sensor domain of XPC

A



B

```
RPA-B (306-411)   –DDLENKSKD–(14)–
XPC (621-730)     –EDLEFQAKH–(5)––
Consensus          nDLE.puKp

                  β1'      β2       β3        β4
RPA-B             –ATKITVRSNNREVAKRNIYLMDTSGKVVTATLWGEDADKFDGSRQPVLA–
XPC               –PTAIGLYKNHPLYALKRHLLKYEAIYPETAAILGYCRGEAVYSRDCVHT–
Consensus          sT.Isa.pNp..hA.+p.hL..pu..s.TAsahG.s.schs.SRpsV.s

                  β4'      β5'     β5
RPA-B             ––––IKGARVSDFG––––RSLSVLSSSTIIANPDIPEAYKL–
XPC               (7)–LKKARVVRLG–(6)–KGFSNRARKARLAEPQLREENDL–
Consensus             aK.ARVschG      +uhSs.upps.aApPpa.E..cL
```

Figure 8. Versatile Damage Recognition: Detection of Single-Stranded Configurations in the Undamaged Strand of the Double Helix

(A) Initial interaction of XPC protein with damaged sites driven by an affinity for native single-stranded DNA. The triangle symbolizes a helix-distorting bulky lesion. This mechanism with inverted DNA strand specificity directs XPC protein to the undamaged strand and the downstream factors of the GGR pathway to the damaged strand.

(B) Alignment of the RPA-B and XPC DNA-binding sequences. The consensus was derived using the following amino acid classes [47]: hydrophobic (h, ALICVMYFW); the aliphatic subset of these (a, ALIVMC); small (s, ACDGNPSTV); the "tiny" subset of these (u, GAS); polar (p, CDEHKNQRST); charged (c, DEHKR), positively charged (+, HKR); and negatively charged (n, DE). The length of nonalignable gaps is indicated in parentheses and the β-sheet elements are indicated by the arrows.

doi:10.1371/journal.pbio.0050079.g008

protein, responsible for the recognition of DNA damage in the GGR pathway, is related to the OB-folds of known single-stranded DNA-binding proteins.

In conclusion, this article shows that a versatile sensor of DNA damage achieves its wide recognition function by avoiding direct contacts with injured residues. Instead, XPC protein exploits the inherent redundancy of the genetic code in the DNA double helix to detect DNA damage in an indirect but highly versatile manner. If one strand contains a bulky lesion, normal base pairing and stacking interactions are compromised, and the intact complementary strand converts to a local single-stranded configuration, thus generating the universal molecular signal for XPC recruitment.

Materials and Methods

Site-directed mutagenesis. The human *XPC* complementary DNA [38] was cloned into pcDNA3.1 (Invitrogen, http://www.invitrogen.com) using the restriction enzymes *Not*I and *Kpn*I and into pEGFP-N3 (Clontech, http://www.clontech.com) using the *Kpn*I and *Xma*I sites. Mutagenesis was carried out with the QuickChange site-directed mutagenesis kit (Stratagene, http://www.stratagene.com) following the manufacturer's instructions. Forward and reverse primers are listed in Table S1. The resulting clones were sequenced (Microsynth, http://www.microsynth.ch) to exclude accidental mutations introduced elsewhere in the complementary DNA.

Host cell reactivation assay. Simian virus 40-transformed human XP–C fibroblasts (GM16093) were from the Coriell Cell Repository (http://ccr.coriell.org). These cells were grown in Dulbecco's modified Eagle's medium (Gibco, http://www.invitrogen.com), supplemented with 10% fetal bovine serum, penicillin G (100 units/ml) and streptomycin (100 μg/ml), in a 5% CO_2 humidified incubator. The pGL3 and phRL–TK vectors expressing firefly (*Photinus*) and *Renilla* luciferase, respectively, were from Promega. DNA was UV-irradiated at a concentration of 1 mg/ml in 10 mM Tris-HCl, (pH 8), and 1 mM EDTA. XP–C cells were transfected in a 6-well plate at a confluence of 95% using Lipofectamine Plus reagent (Invitrogen). Each transfection mixture contained 0.23 μg pGL3 (UV-irradiated), 0.02 μg phRL–TK (unirradiated), and 0.25 μg of the appropriate expression vector. After a 4-h incubation, the transfection reagents were replaced by complete medium. Unless otherwise indicated, the cells were lysed after another 15-h period to measure firefly and *Renilla* luciferase activity using the Dual–Luciferase assay system (Promega, http://www.promega.com) on a microtiter plate luminometer (Dynex, http://www.dynextechnologies.com). All results (mean values of at least five determinations) were normalized by calculating the ratios between firefly and *Renilla* luciferase activity. Expression of XPC polypeptides in human cells was monitored by Western blotting (using monoclonal antibodies against GFP from Clontech) and fluorescence microscopy as described [46].

Expression and purification of XPC protein. A polyhistidine-MBP–XPC fusion product was constructed by inserting a 2.9-kb fragment, which contains the human *XPC* complementary DNA, into the pFastBac HTc vector (Invitrogen) using the *Not*I and *Kpn*I restriction sites. Subsequently, a 1.2-kb fragment containing the *MalE* complementary DNA (from pMal-c2; New England Biolabs, http://www.neb.com) was inserted on the 5′ side of the *XPC* sequence using the *Stu*I restriction site. Recombinant baculovirus for the infection of *Sf*9 cells was generated using the BAC-TO-BAC Baculovirus Expression System (Invitrogen) following the manufacturer's instructions.

Polyhistidine- and MBP-tagged XPC protein was fractionated from *Sf*9 cell lysates [38] with two chromatographic cycles through a Ni^{2+} column (Qiagen, http://www.qiagen.com). The fractions were analyzed using a mouse monoclonal antibody against recombinant human XPC protein (Abcam, http://www.abcam.com). Samples containing XPC protein, eluting mainly at 100 mM imidazole, were pooled, dialyzed against phosphate buffer (25 mM sodium phosphate, [pH 7.8], 10% [v/v] glycerol, 5 mM β-mercaptoethanol, and 0.25 mM phenylmethane sulfonyl fluoride) containing 0.2 M NaCl, and further processed by heparin chromatography (Amersham, http://www.amershambiosciences.com). The heparin column was eluted with a 0.2–1 M gradient of NaCl. The samples containing homogeneous MBP–XPC protein, eluting at 600 mM NaCl, were pooled, dialyzed, and supplemented with glycerol to a concentration of 25% (v/v) before freezing at −80 °C. Protein concentration was determined using the Bio-Rad protein assay reagent (http://www.bio-rad.com).

A one-step purification was performed by mixing crude *Sf*9 cell lysates (5–20 μl) with monoclonal antibodies against MBP that were covalently linked to paramagnetic beads (New England BioLabs). The binding buffer consisted of 25 mM Tris-HCl, (pH 7.5), 10% glycerol, 0.01% Triton X-100, 0.25 mM phenylmethane sulfonyl fluoride, 1 mM EDTA, and 0.3 M NaCl. After incubation at 4 °C for 2 h, the beads were washed four times, and bound proteins were analyzed by denaturing gel electrophoresis followed by Coomassie staining. The yield of MBP–XPC protein was determined by quantitative laser densitometry of the 170-kDa bands using, as standards, different amounts of MBP–XPC probes purified by Ni^{2+} and heparin chromatography, as described before, and loaded in parallel onto the same gel.

Electrophoretic mobility shift assays. The synthetic 65-mer oligonucleotides 5′-CGGGGCGAATTCGAGCTCGCCCGGGATCCTCACATAGAGTCGACCTGCTGCAGCCCAAGCTTGGC-3′ and 5′-GCCAAGCTTGGGCTGCAGCAGGTCGACTCTATGTGAGGATCCCGGGCGAGCTCGAATTCGCCCCG-3′ were purchased from Microsynth. A DNA homoduplex was constructed by hybridizing these complementary oligonucleotides in 50 mM Tris-HCl, (pH 7.4), 10 mM $MgCl_2$, and 1 mM dithiothreitol. The annealing was performed by heating to 95 °C for 10 min, followed by slow cooling (3 h at 25 °C). Electrophoretic mobility shift assays (reactions of 10 μl) were performed by incubating, at 20 °C for 30 min, ^{32}P-labeled oligonucleotide substrate (2 nM), duplex poly[dI–dC] competitor DNA (10 ng/μl), and the indicated concentrations of XPC protein in

40 mM Tris-HCl, (pH 7.5), 5 mM $MgCl_2$, 100 μg/ml bovine serum albumin, and 1 mM dithiothreitol [40]. Following the addition of gel loading buffer (2 μl) containing 30% (v/v) glycerol, 0.25% (w/v) bromophenol blue, and 0.25% (w/v) xylene cyanol in water, the extent of binding was determined on 7% native polyacrylamide gels.

Screening of mutants for DNA binding. Lysates (5 μl) from baculovirus-infected *Sf*9 cells [28] were mixed with 50 μl of single-stranded DNA agarose beads (Amersham) and 100 μl 25 mM Tris-HCl, (pH 7.5), 10% glycerol, 0.01% Triton X-100, 0.25 mM phenylmethane sulfonyl fluoride, 1 mM EDTA, supplemented by the indicated concentrations of NaCl. After incubation at 4 °C for 2 h, the supernatant was recovered and the beads were washed four times with 300 μl binding buffer. Finally, the DNA-bound proteins were eluted from the beads with 100 μl of 10 mM Tris-HCl, (pH 8.0), 1 mM EDTA, and 1% (w/v) sodium dodecylsulfate. Equivalent amounts of supernatant and DNA-bound fractions were loaded onto denaturing polyacrylamide gels, followed by immunoblot analysis, visualization by chemoluminescence (SuperSignal, Pierce, http://www.piercenet.com), and quantification by laser scanning densitometry.

The binding of mutants to single-stranded or double-stranded oligonucleotides was tested using purified MBP–XPC fusions obtained by immunoprecipitation. Paramagnetic beads (0.2 mg) containing the indicated amounts of wild-type or mutant XPC (between 10 and 100 ng) were incubated with ^{32}P-labeled 65-mer probes (2 nM) in 200 μl of 25 mM Tris-HCl, (pH 7.5), 0.3 M NaCl, 10% glycerol, 0.01% Triton X-100, 0.25 mM phenylmethane sulfonyl fluoride, and 1 mM EDTA. Following an incubation of 90 min at 4 °C, the paramagnetic beads were washed three times with 200-μl binding buffer. Finally, the radiolabeled oligonucleotides associated with XPC protein were quantified by liquid scintillation counting. The background radioactivity resulting from unspecific binding of the oligonucleotides to empty beads (0.2 mg) was determined in separate reactions.

Supporting Information

Table S1. Forward and Reverse Primers Used for Site-Directed Mutagenesis

The mutated nucleotides are underlined.

Found at doi:10.1371/journal.pbio.0050079.st001 (62 KB PDF).

Acknowledgments

We thank M. Träxler and M. Vitanescu for excellent technical assistance and L. Staresincic for advice in XPC purification. We are also grateful to J. T. Reardon and A. Sancar for the human *XPC* complementary DNA.

Author contributions. OM, SS, and HN conceived and designed the experiments. OM and SS performed the experiments and analyzed the data. HN wrote the paper.

Funding. This work is supported by Swiss National Science Foundation grant 3100A–113694.

Competing interests. The authors have declared that no competing interests exist.

References

1. Sancar A (1996) DNA excision repair. Annu Rev Biochem 65: 43–81.
2. Wood RD (1997) Nucleotide excision repair in mammalian cells. J Biol Chem 272: 23465–23468.
3. De Laat WL, Jaspers NG, Hoeijmakers JHJ (1999) Molecular mechanism of nucleotide excision repair. Genes Dev 13: 768–785.
4. Gillet LC, Schärer OD (2006) Molecular mechanisms of mammalian global genome nucleotide excision repair. Chem Rev 106: 253–276.
5. Satoh MS, Jones CJ, Wood RD, Lindahl T (1993) DNA excision-repair defect of xeroderma pigmentosum prevents removal of a class of oxygen free radical-induced base lesions. Proc Natl Acad Sci U S A 90: 6335–6339.
6. Reardon JT, Bessho T, Kung HC, Bolton PH, Sancar A (1997) In vitro repair of oxidative DNA damage by human nucleotide excision repair system: Possible explanation for neurodegeneration in xeroderma pigmentosum patients. Proc Natl Acad Sci U S A 94: 9463–9468.
7. Kuraoka I, Bender C, Romieu A, Cadet J, Wood RD, Lindahl T (2000) Removal of oxygen free-radical-induced 5′,8-purine cyclodeoxynucleosides from DNA by the nucleotide excision-repair pathway in human cells. Proc Natl Acad Sci USA 97: 3832–3837.
8. Reardon JT, Sancar A (2006) Repair of DNA-polypeptide crosslinks by human excision nuclease. Proc Natl Acad Sci U S A 103: 4056–4061.
9. Cleaver JE, Thompson LH, Richardson AS, States JC (1999) A summary of mutations in the UV–sensitive disorders: Xeroderma pigementosum, Cockayne syndrome, and trichothiodystrophy. Hum Mutat 14: 9–22.
10. Friedberg EC (2001) How nucleotide excision repair protects against cancer. Nat Rev Cancer 1: 22–33.
11. Andresso JO, Hoeijmakers JHJ (2005) Transcription-coupled repair and premature ageing. Mutat Res 577: 179–194.
12. Cleaver JE (2005) Cancer in xeroderma pigmentosum and related disorders of DNA repair. Nat Rev Cancer 5: 564–573.
13. Kraemer KH, Lee MM, Scotto J (1984) DNA repair protects against cutaneous and internal neoplasia: Evidence from xeroderma pigmentosum. Carcinogenesis 5: 511–514.
14. Kraemer KH, Lee MM, Scotto J (1987) Xeroderma pigmentosum. Cutaneous, ocular, and neurologic abnormalities in 830 published cases. Arch Dermatol 123: 241–250.
15. Friedberg EC, Walker GC, Siede W (1995) DNA repair and mutagenesis. Washington, D. C.: ASM Press. 698 p.
16. Hanawalt PC (2002) Subpathways of nucleotide excision repair and their regulation. Oncogene 21: 8949–8956.
17. Mellon I (2005) Transcription-coupled repair: A complex affair. Mutat Res 577: 155–161.
18. Lainé JP, Egly J-M (2006) Initiation of DNA repair mediated by a stalled RNA polymerase IIO. EMBO J 25: 387–397.
19. Sugasawa K, Ng JM, Masutani C, Iwai S, van der Spek PJ, et al. (1998) Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. Mol Cell 2: 223–232.
20. Volker M, Moné MJ, Karmakar P, van Hoffen A, Schul W, et al. (2001) Sequential assembly of the nucleotide excision repair factors In vivo. Mol Cell 8: 213–224.
21. Bergink S, Salomons FA, Hoogstraten D, Groothuis TAM, de Waard H, et al. (2006) DNA damage triggers nucleotide excision repair-dependent monoubiquitylation of histone H2A. Genes Dev 20: 1343–1352.
22. Marti TM, Hefner E, Feeney L, Natale V, Cleaver JE (2006) H2AX phosphorylation within the G1 phase after UV irradiation depends on nucleotide excision repair and not DNA double-strand breaks. Proc Natl Acad Sci U S A 103: 9891–9896.
23. Sugasawa K, Shimizu Y, Iwai S, Hanaoka F (2002) A molecular mechanism for DNA damage recognition by the xeroderma pigmentosum group C protein complex. DNA Rep 1: 95–107.
24. Buterin T, Meyer C, Giese B, Naegeli H (2005) DNA quality control by conformational readout on the undamaged strand of the double helix. Chem Biol 12: 913–922.
25. Trego KS, Turchi JJ (2006) Pre-steady-state binding of damaged DNA by XPC–hHR23B reveals a kinetic mechanism for damage discrimination. Biochemistry 45: 1961–1969.
26. Sugasawa K, Masutani C, Uchida A, Maekawa T, van der Spek PJ, et al. (1996) hHR23B, a human Rad23 homolog, stimulates XPC protein in nucleotide excision repair in vitro. Mol Cell Biol 16: 4852–4861.
27. Nishi R, Okuda Y, Watanabe E, Mori T, Iwai S, et al. (2005) Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein. Mol Cell Biol 25: 5664–6574.
28. Uchida A, Sugasawa K, Masutani C, Dohmae N, Araki M, et al. (2002) The carboxy–terminal domain of the XPC protein plays a crucial role in nucleotide excision repair through interactions with transcription factor IIH. DNA Rep 1: 449–461.
29. Chavanne F, Broughton BC, Pietra D, Nardo T, Browitt A, et al. (2000) Mutations in the XPC gene in families with xeroderma pigmentosum and consequences at the cell, protein, and transcript level. Cancer Res 60: 1974–1982.
30. Murzin AG (1993) OB(oligonucleotide/oligosaccharide binding)-fold: Common structural and functional solution for non-homologous sequences. EMBO J 12: 861–867.
31. Bochkarev A, Pfuetzner RA, Edwards AM, Frappier L (1997) Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. Nature 385: 176–181.
32. Theobald DL, Mitton-Fry RM, Wuttke DS (2003) Nucleic acid recognition by OB-fold proteins. Annu Rev Biophys Biomol Struct 32: 115–133.
33. Bochkarev A, Bochkareva E (2004) From RPA to BRCA2: Lessons from single-stranded DNA binding by the OB-fold. Curr Opin Struct Biol 14: 36–42.
34. Notredame C, Higgins DG, Heringa J (2000) T–Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol 302: 205–217.
35. Carreau M, Eveno E, Quilliet X, Chevalier–Lagente O, Benoit A, et al. (1995) Development of a new easy complementation assay for DNA repair deficient human syndromes using cloned repair genes. Carcinogenesis 16: 1003–1009.
36. Kapust RB, Waugh DS (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. Protein Sci 8: 1668–1674.
37. Alvarez D, Callejo M, Shoucri R, Boyer L, Price GB, et al. (2003) Analyis of the cruciform binding activity of recombinant 14-3-3zeta-MBP fusion protein, its heterodimerization profile with endogenous 14-3-3 isoforms,

and effect on mammalian DNA replication in vitro. Biochemistry 42: 7205–7215.
38. Reardon JT, Mu D, Sancar A (1996) Overproduction, purification, and characterization of the XPC subunit of the human repair excision nuclease. J Biol Chem 271: 19451–19456.
39. Hey T, Lipps G, Sugasawa K, Iwai S, Hanaoka F, et al. (2002) The XPC-HR23B complex displays high affinity and specificity for damaged DNA. Biochemistry 41: 6583–6587.
40. Batty D, Rapic'-Otrin V, Levine AS, Wood RD (2000) Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged sites. J Mol Biol 300: 275–290.
41. Sugasawa K, Okamoto T, Shimizu Y, Masutani C, Iwai S, et al. (2001) A multistep damage recognition mechanism for global genomic nucleotide excision repair. Genes Dev 15: 507–521.
42. Liu Y, Yang Z, Utzat CD, Liu Y, Geacintov NE, et al. (2005) Interactions of human replication protein A with single-stranded DNA adducts. Biochem J 385: 519–526.
43. Arcus V (2002) OB-fold domains: A snapshot of the evolution of sequence, structure and function. Curr Opin Struct Biol 12: 794–801.
44. Legerski RJ, Peterson C (1992) Expression cloning of a human DNA repair gene involved in xeroderma pigmentosum group C. Nature 359: 70–73.
45. Henning KA, Peterson C, Legerski R, Friedberg EC (1994) Cloning of the *Drosophila* homolog of the xeroderma pigmentosum complementation group C gene reveals homology between the predicted human and *Drosophila* polypeptides and that encoded by the yeast *RAD4* gene. Nucleic Acids Res 22: 257–261.
46. Ng JMY, Vermeulen W, van der Horst GTJ, Bergink S, Sugasawa K, et al. (2003) A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. Genes Dev 17: 1630–1645.
47. Anantharaman V, Koonin EV, Aravind L (2001) Peptide-*N*-glycanases and DNA repair proteins, Xp-C/Rad4, are, respectively, active and inactivated enzymes sharing a common transglutaminase fold. Hum Mol Genet 10: 1627–1630.

Table 1

Forward and reverse primers used for site-directed mutagenesis. The mutated nucleotides are underlined.

Substitution	Oligonucleotide primer sequences
W531G	5'-GCTGGTATAGACCAG**G**GGCTAGAGGTGTTC-3' 5'-GAACACCTCTAGCC**C**CTGGTCTATACCAGC-3'
W531A	5'-GCTGGTATAGACCAGG**C**GCTAGAGGTGTTC-3' 5'-GAACACCTCTAGC**G**CCTGGTCTATACCAGC-3'
W531S	5'-GCTGGTATAGACCAGT**C**GCTAGAGGTGTTC-3' 5'-GAACACCTCTAGC**G**ACTGGTCTATACCAGC-3'
F535A	5'-CAGTGGCTAGAGGTG**GC**CTGTGAGCAGGAG-3' 5'-CTCCTGCTCACAG**GC**CACCTCTAGCCACTG-3'
W542G	5'-GAGCAGGAGGAAAAG**G**GGGTATGTGTAGAC-3' 5'-CTCCTCTGTCTGCC**C**CTTTTCCTCCTGCTC-3'
W542A	5'-GAGCAGGAGGAAAAGG**C**GGTATGTGTAGAC-3' 5'-CTCCTCTGTCTGCG**C**CTTTTCCTCCTGCTC-3'
W542S	5'-GAGCAGGAGGAAAAGT**C**GGTATGTGTAGAC-3' 5'-CTCCTCTGTCTGC**G**ACTTTTCCTCCTGCTC-3'
Y585A	5'-GATGTCACACAGAGG**GC**CGACCCAGTCTGG-3' 5'-CCAGACTGGGTCG**GC**CCTCTGTGTGACATC-3'
Y585G	5'-GATGTCACACAGAGGG**G**CGACCCAGTCTGG-3' 5'-CCAGACTGGGTCG**C**CCCTCTGTGTGACATC-3'
Y585S	5'-GATGTCACACAGAGG**T**CCGACCCAGTCTGG-3' 5'-CCAGACTGGGTCGG**A**CCTCTGTGTGACATC-3'
K595G	5'-GGATGACAGTGACCCGC**GG**GTGCCGGGTTGATGCTG-3' 5'-CAGCATCAACCCGGCAC**CC**GCGGGTCACTGTCATCC-3'
R597G	5'-CAGTGACCCGCAAGTGC**G**GGGTTGATGCTGAGTGG-3' 5'-CCACTCAGCATCAACCC**C**GCACTTGCGGGTCACTG-3'
P635A	5'-CACATGGACCAGCCTTTG**G**CCACTGCCATTGGC-3' 5'-GCCAATGGCAGTGG**C**CAAAGGCTGGTCCATGTG-3'
Y641A	5'-CCCACTGCCATTGGCTTA**GC**TAAGAACCACCCTCTG-3' 5'-CAGAGGGTGGTTCTTA**GC**TAAGCCAATGGCAGTGGG-3'
K642G	5'-CTGCCATTGGCTTATAT**GG**GAACCACCCTCTGTATGC-3' 5'-GCATACAGAGGGTGGTTC**CC**ATATAAGCCAATGGCAG-3'
H644A	5'-GGCTTATATAAGAAC**GC**CCCTCTGTATGCCCTG-3' 5'-CAGGGCATACAGAGGG**GC**GTTCTTATATAAGCC-3'
Y647A	5'-GAACCACCCTCTG**GC**TGCCCTGAAGCGGCATCTC-3' 5'-GAGATGCCGCTTCAGGGCA**GC**CAGAGGGTGGTTC-3'
R651G	5'-CCTCTGTATGCCCTGAAG**G**GGCATCTCCTGAAATATG-3' 5'-CATATTTCAGGAGATGCC**C**CTTCAGGGCATACAGAGG-3'

K655G	5'-GAAGCGGCATCTCCTG**GG**ATATGAGGCCATCTATCC-3' 5'-GGATAGATGGCCTCATAT**CC**CAGGAGATGCCGCTTC-3'
Y656A	5'-CGGCATCTCCTGAAA**GC**TGAGGCCATCTATCCC-3' 5'-GGGATAGATGGCCTCA**GC**TTTCAGGAGATGCCG-3'
Y660A	5'-GAAATATGAGGCCATC**GC**TCCCGAGACAGCTGCC-3' 5'-GGCAGCTGTCTCGGGA**GC**GATGGCCTCATATTTC-3'
Y676A	5'-CGTGGAGAAGCGGTC**GC**CTCCAGGGATTGTGTG-3' 5'-CACACAATCCCTGGAG**GC**GACCGCTTCTCCACG-3'
R678G	5'-GAAGCGGTCTACTCC**G**GGGATTGTGTGCACAC-3' 5'-GTGTGCACACAATCCC**C**GGAGTAGACCGCTTC-3'
S686A	5'-GTGCACACTCTGCAT**G**CCAGGGACACGTGGCTG-3' 5'-CAGCCACGTGTCCCTGG**C**ATGCAGAGTGTGCAC-3'
W690G	5'-GCATTCCAGGGACACG**G**GGCTGAAGAAAGCAAGAG-3' 5'-CTCTTGCTTTCTTCAGCC**C**CGTGTCCCTGGAATGC-3'
W690A	5'-GCATTCCAGGGACACG**G**CGCTGAAGAAAGCAAGAG-3' 5'-CTCTTGCTTTCTTCAGCG**C**CGTGTCCCTGGAATGC-3'
W690S	5'-GCATTCCAGGGACACGT**C**GCTGAAGAAAGCAAGAG-3' 5'-CTCTTGCTTTCTTCAGC**G**ACGTGTCCCTGGAATGC-3'
K693G	5'-GGACACGTGGCTGAAG**GG**AGCAAGAGTGGTGAGG-3' 5'-CCTCACCACTCTTGCT**CC**CTTCAGCCACGTGTCC-3'
R698G	5'-GAAAGCAAGAGTGGTG**G**GGCTTGGAGAAGTACC-3' 5'-GGTACTTCTCCAAGCC**C**CACCACTCTTGCTTTC-3'
Y704A	5'-CTTGGAGAAGTACCC**GC**CAAGATGGTGAAAGGC-3' 5'-GCCTTTCACCATCTTG**GC**GGGTACTTCTCCAAG-3'
K705G	5'-GCTTGGAGAAGTACCCTAC**GG**GATGGTGAAAGGCTTTTC-3' 5'-GAAAAGCCTTTCACCATC**CC**GTAGGGTACTTCTCCAAGC-3'
K708G	5'-CCCTACAAGATGGTG**GG**AGGCTTTTCTAACCGTGC-3' 5'-GCACGGTTAGAAAAGCCT**CC**CACCATCTTGTAGGG-3'
F733A	5'-GAAAATGACCTGGGCCTG**GC**TGGCTACTGGCAGACAGAG-3' 5'-CTCTGTCTGCCAGTAGCCA**GC**CAGGCCCAGGTCATTTTC-3'
F733G	5'-GAAAATGACCTGGGCCTGG**G**TGGCTACTGGCAGACAGAG-3' 5'-CTCTGTCTGCCAGTAGCCA**C**CCAGGCCCAGGTCATTTTC-3'
F733S	5'-GAAAATGACCTGGGCCTG**T**CTGGCTACTGGCAGACAGAG-3' 5'-CTCTGTCTGCCAGTAGCCAG**A**CAGGCCCAGGTCATTTTC-3'
W736G ?	5'-GGCCTGTTTGGCTAC**G**GGCAGACAGAGGAG-3' 5'-CTCCTCTGTCTGCC**C**GTAGCCAAACAGGCC-3'
Y741A	5'-GGCAGACAGAGGAG**GC**TCAGCCCCCAGTG-3' 5'-CACTGGGGGCTGA**GC**CTCCTCTGTCTGCC-3'
F756A	5'-GTGCCCCGGAACGAG**GC**TGGGAATGTGTAC-3' 5'-GTACACATTCCCA**GC**CTCGTTCCGGGGCAC-3'
F762A	5'-GGGAATGTGTACCTC**GC**CCTGCCCAGCATG-3' 5'-CATGCTGGGCAGG**GC**GAGGTACACATTCCC-3'
F762G	5'-GGGAATGTGTACCTCG**G**CCTGCCCAGCATG-3' 5'-CATGCTGGGCAGG**C**CGAGGTACACATTCCC-3'
F762S	5'-GGGAATGTGTACCTC**T**CCCTGCCCAGCATG-3' 5'-CATGCTGGGCAGGG**A**GAGGTACACATTCCC-3'

F797A	5’-CAGGCCATCACTGGC**GC**TGATTTCCATGGCGG-3’ 5’-CCGCCATGGAAATCA**GC**GCCAGTGATGGCCTG-3’
F799A	5’-CCATCACTGGCTTTGAT**GC**CCATGGCGGCTACTC-3’ 5’-GAGTAGCCGCCATGG**GC**ATCAAAGCCAGTGATGG-3’

Chapter III

Versatile Protection from Mutagenic DNA Lesions Conferred by Bipartite Recognition in Nucleotide Excision repair

Versatile protection from mutagenic DNA lesions conferred by bipartite recognition in nucleotide excision repair

Olivier Maillard[1], Ulrike Camenisch[1], Krastan B. Blagoev[2] and Hanspeter Naegeli[1]

[1]Institute of Pharmacology and Toxicology, University of Zürich-Vetsuisse, Winterthurerstrasse 260, CH-8057 Zürich, Switzerland

[2]Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

Abstract

Nucleotide excision repair is a cut-and-patch pathway that eliminates potentially mutagenic DNA lesions caused by ultraviolet light, electrophilic chemicals, oxygen radicals and many other genetic insults. Unlike antigen recognition by the immune system, which employs billions of immunoglobulins and T-cell receptors, the nucleotide excision repair complex relies on just a few generic factors to detect an infinite range of DNA adducts. This molecular versatility is achieved by a bipartite strategy initiated by the detection of abnormal DNA fluctuations, followed by the localization of injured residues through an enzymatic scanning process. The early recognition subunits are able to probe the dynamic properties of nucleic acid substrates but avoid direct contacts with chemically altered bases. Only downstream subunits of the bipartite recognition process interact more closely with damaged bases to delineate the sites of DNA incision. Thus, consecutive factors expand the spectrum of deleterious genetic lesions conveyed to DNA repair by detecting distinct molecular features of target substrates.

Keywords. Mutagenesis, cancer, carcinogens, genome stability, DNA repair, ultraviolet radiation, xeroderma pigmentosum

1. Introduction

Although integrity of the genetic information is essential for normal development, viability, longevity and the health of organisms, the cellular DNA is under permanent attack not only from environmental genotoxic agents but also from endogenous metabolic byproducts that alter its chemical structure. To counteract the continuous formation of genetic damage, living organisms are equipped with a network of DNA repair systems. Briefly, placental mammals employ six major DNA repair pathways to cope with mutagenic insults [1]: (i) mismatch repair to correct replication errors, (ii) DNA damage reversal to remove alkyl groups, (iii) non-homologous end joining to repair double-strand breaks, (iv) homologous recombination to rescue corrupted or deleted chromosomal sequences, (v) base excision repair to eliminate modified or incorrect bases, and (vi) nucleotide excision repair (NER) to remove bulky lesions.

The NER system eliminates DNA lesions by promoting the excision of single-stranded oligonucleotides from damaged strands followed by restoration of an intact double helix by DNA repair synthesis and DNA ligation (Fig. 1). This type of repair reaction has evolved in all three biological kingdoms to excise mutagenic photoproducts induced by short-wavelength ultraviolet (UV) light [primarily cyclobutane pyrimidine dimers and (6-4) photoproducts] as well as a wide array of bulky DNA adducts generated by electrophilic carcinogens [2-4]. Other known NER substrates include a subset of oxidative lesions [5-7] and protein-DNA crosslinks [8]. The purpose of this review is to summarize the knowledge gained in the last few years as to how the mammalian NER system achieves this astounding substrate versatility while preserving a stringent selectivity for damaged target sequences.

The importance of preventing genetic mutations caused by DNA photoproducts and other NER substrates is illustrated by a direct link between defects in the NER pathway and a devastating cancer-prone disorder in humans. In fact, many NER proteins are encoded by

genes that, when mutated, give rise to xeroderma pigmentosum (XP), an inherited disease characterized by extreme photosensitivity and a 2000-fold increased incidence of sunlight-induced skin cancer [1,9]. XP patients also have a higher risk of internal tumors and, in some cases, neurological complications, probably reflecting the essential role of NER in the removal of oxidative DNA damage [5,10]. Individuals suffering from this recessive disorder have been assigned to different complementation groups by cell-fusion experiments and the respective NER genes (*XPA* through *XPG*) were named after the complementation group with which they associate.

2. Mammalian NER reaction and core subunits

The *XPA-XPG* gene products are the components of a NER complex that is necessary and sufficient for excision activity in the presence of naked substrates *in vitro* without the aid of any other accessory protein [11-13]. The order of arrival of these core NER factors is still debated, but a favored model illustrated in Fig. 1 proposes the assembly of a multi-subunit complex triggered by XPC protein together with one of the mammalian homologs of yeast RAD23 (HR23B) and the calcium-binding protein centrin-2 (CTN2) [14-16]. After initial recognition of damaged sites by the XPC subunit, this pathway is thought to proceed with the sequential recruitment of transcription factor IIH (TFIIH, containing XPB, XPD and other 8 subunits), XPA (a possible homodimer), replication protein A (RPA, 3 subunits), XPG and XPF-ERCC1 (a dimer composed of XPF and excision repair cross complementing-1 protein).

Irrespective of the order of NER assembly, which will be discussed in more detail below, it has been demonstrated that XPA, XPC, TFIIH and RPA participate in the formation of a stable recognition intermediate [17,18] characterized by transient unwinding of the duplex substrate (Fig. 1). In conjunction, these factors introduce an open DNA structure with "Y-shaped" double to single strand junctions flanking the lesion [18-20]. The endonucleases

XPF-ERCC1 and XPG act as “scissors” to cut out DNA damage by cleaving the damaged strand at each of the “Y-shaped” transitions of this open intermediate, thereby releasing injured residues as part of an oligomeric segments of 24-32 nucleotides in length [21]. XPF-ERCC1 makes the 5’ incision, whereas XPG is responsible for the 3’ incision [22-24]. The DNA scissions are introduced 15-25 nucleotides away from the damaged base on the 5’ side but only 3-9 nucleotides away on the 3’ side.

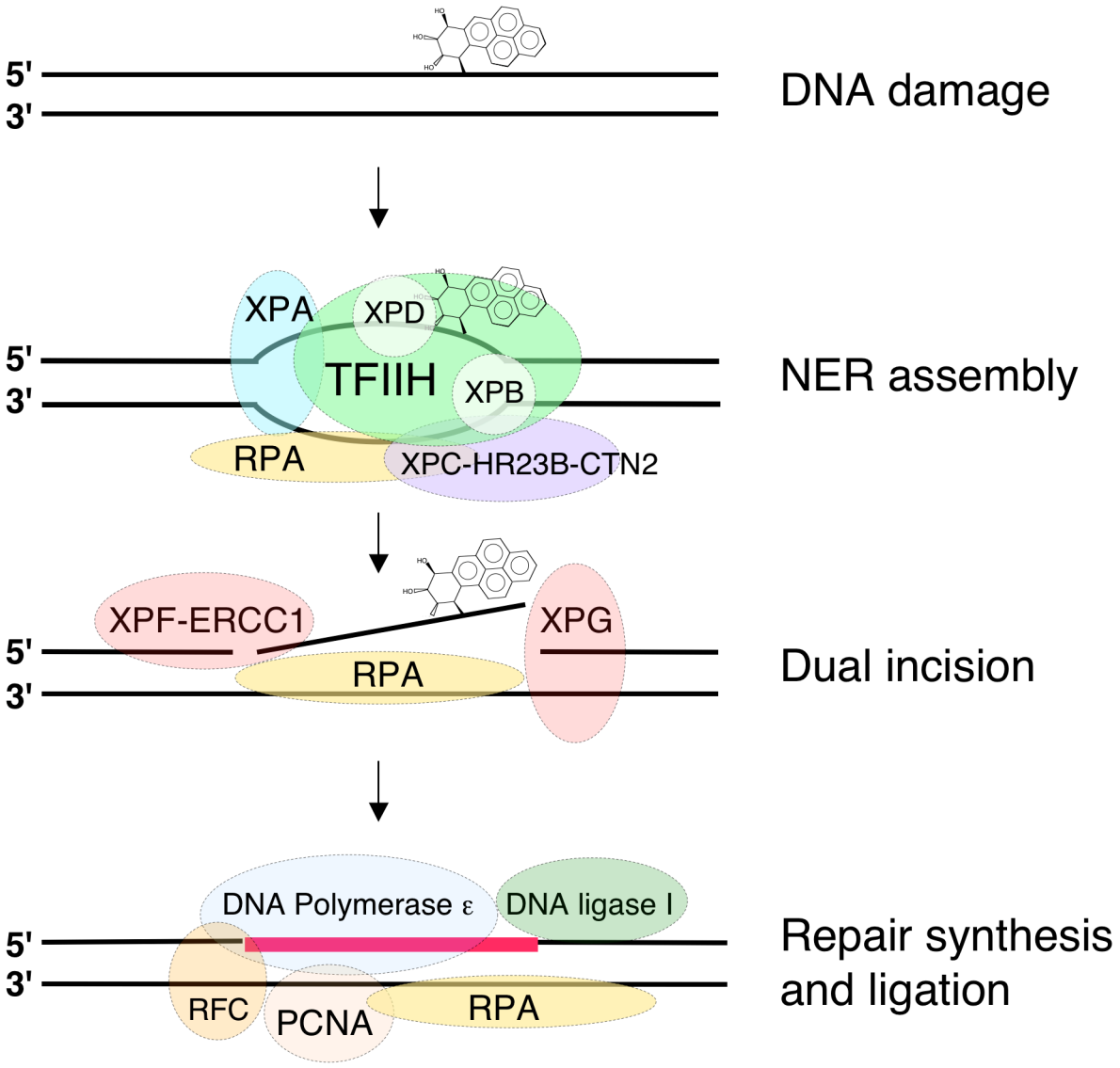


Figure 1. Scheme of the human NER pathway. The damaged strand carries a polycyclic aromatic hydrocarbon adduct. Abbreviations: ERCC1, excision repair cross complementing-1 protein; HR23B, homolog of RAD23B; PCNA, proliferating cell nuclear antigen; RFC, replication factor C; RPA, replication protein A; TFIIH, transcription factor IIH; XPA-XPF, xeroderma pigmentosum complementation group A-F proteins. DNA damage recognition by the action of XPC, XPA, TFIIH and RPA is followed by double DNA incision through the endonucleases XPF-ERCC1 and XPG. The synthesis of repair patches is dependent on RFC, a matchmaker that binds to the excision gap and loads PCNA, which in turn acts as a sliding clamp for DNA polymerases. Finally, the newly synthesized repair patches are sealed by DNA ligase I.

To restore duplex integrity, all excision repair reactions depend on the redundancy of the double-stranded genetic code. If the nucleotides of one strand are damaged, they are excised and the intact opposing strand is used as a template to direct the synthesis of repair patches. Because the NER pathway generates two separate incisions, it is essential that both scissions occur in the same damaged strand, such that the opposite native sequence is preserved during the excision reaction and later in the pathway can serve as the complementary template for error-free DNA synthesis [11,25].

3. Assembly of the DNA damage recognition complex

The mechanism of DNA lesion recognition in the NER pathway is a matter of intense debate, raised by the fact than none of the individual core proteins displays a high enough specificity to function as a unique sensor of damaged substrates. In mammals, the NER reaction occurs by the individual recruitment of repair factors to sites of damage, rather than by the action of a preassembled “repairosome”. A full excision complex with approximately 20 polypeptide

subunits would achieve a mass of >1 MDa, but studies monitoring diffusion rates in the nuclei of intact cells indicate that the NER proteins are present as separate factors and not as part of a large "repairosome" complex. Also, the rapid translocation of NER factors to foci of DNA damage is not compatible with the existence of large preassembled complexes and instead favors the individual recruitment of each subunit to lesion sites [26]. One advantage of this stepwise process is that multifunctional proteins may shuttle between repair, transcription, replication, recombination or other nuclear pathways [27]. In addition, the mathematical modeling of various scenarios suggests that an ordered and consecutive assembly of freely diffusing proteins is more efficient than alternative strategies such as the random aggregation of repair factors, or their preassembly into a "repairosome" complex [28].

Initially, two opposing mechanisms have been proposed for the damage recognition step: "XPC first" or "XPA first" [18]. In the "XPC first" model, XPC represents the primary sensor that binds to lesion sites and initiates the NER pathway by recruiting TFIIH and other successive factors [29-31]. This scenario is supported by competition experiments aimed at determining the order in which NER proteins are recruited to the DNA substrate. For example, damaged plasmids preincubated with XPC-HR23B are more rapidly repaired in cell extracts than those previously incubated with XPA and RPA [30], indicating that the NER complex works more efficiently *in vitro* when XPC is allowed time to bind to DNA before addition of the remaining subunits. In apparent conflict with these reports, other researchers observed that preincubation of damaged DNA with XPA-RPA promotes excision more effectively than when the damaged substrate is first incubated with XPC-HR23B [32]. These conflicting models have been reconciled by the notion that XPC, XPA and RPA may act in a cooperative manner to locate the lesions and recruit the TFIIH complex [33]. A conceptual advantage of this concerted action is that three subunits in conjunction achieve an increased

affinity for damaged DNA by combining the modest selectivity of each component alone. This cooperative model of damage recognition provides a plausible explanation for the efficient *in vitro* excision of cyclobutane dimers, although no single subunit displays a significant affinity for this particular UV lesion [31,34].

Site-directed crosslinking probes have been used to test whether any of the putative DNA damage recognition proteins may interact with injured bases. For that purpose, a reactive furan-side psoralen adduct was constructed to serve both as a substrate for *in vitro* excision assays and as a crosslinking reagent that immobilizes repair subunits located in close proximity to the lesion [35]. When the psoralen adduct was incubated with the whole reconstituted NER system, two subunits of RPA (RPA70 and RPA32) and a single TFIIH subunit (XPD) were crosslinked to the DNA substrate. Surprisingly, neither XPC nor XPA were immobilized to the psoralen moiety, indicating that these factors do not make intimate contacts with the defective nucleotide, even though they bind preferentially to damaged duplexes. In contrast, XPD appears to contribute directly to damage recognition [35].

Another experimental strategy to study NER assembly exploits the nuclear trafficking of core subunits in living cells. Fibroblast monolayers were exposed to UV light through filters with small pores to obtain localized foci of DNA damage. The translocation of XPC and XPA from unirradiated nuclear regions to these damaged foci was monitored by immunological staining after formaldehyde fixation. Volker et al. [14] observed that XPC protein readily accumulates in the DNA repair foci of both wild-type and XP-A cells, but they did not detect any accumulation of XPA protein in the DNA damage foci of XP-C cells. These results have been taken as further evidence for XPC being the first factor that recognizes DNA lesions, whereas XPA is apparently not able to interact with damaged sites in the absence of the XPC subunit.

4. The accessory role of UV-DDB

UV-damaged DNA-binding (UV-DDB) protein has been isolated by virtue of its ability to interact preferentially with UV-irradiated DNA fragments [36]. UV-DDB accelerates the excision of UV dimers in human cells [37], but contradictory results have been reported as to how this factor contributes to the NER reaction [18,38,39]. UV-DDB is a heterodimer of p127 (DDB1) and p48 (DDB2), with the small subunit being encoded by the *XPE* gene [40-42].

UV-DDB is considered to be an initial damage sensor because of its extraordinary preference for UV-irradiated substrates [43,44]. It binds with high affinity to (6-4) photoproducts and, unlike XPA or XPC, also interacts preferentially with DNA duplexes containing cyclobutane pyrimidine dimers [45]. The binding of UV-DDB to damaged substrates leads to bending of the DNA by an angle of 55° [46], prompting the hypothesis that UV-DDB may recognize photoproducts and distort the DNA around the lesion to mediate the subsequent recruitment of XPA or XPC. XP-E cells lacking UV-DDB activity are compromised in the repair of cyclobutane dimers but retain the ability to excise (6-4) photoproducts [47]. Similarly, rodent cells, which fail to express DDB2 protein because of promoter methylation, are inefficient in cyclobutane dimer repair [48]. Finally, DDB2 accumulates at sites of UV-induced DNA lesions and the recruitment of XPC to DNA repair foci containing exclusively cyclobutane dimers has been reported to be dependent on UV-DDB [39,49].

In vitro binding studies revealed that UV-DDB exhibits not only a high selectivity for UV-irradiated DNA but also a moderate affinity for several other lesions [46,50,51]. Electrophoretic mobility shift assays demonstrated that UV-DDB binds most efficiently to DNA fragments containing a (6-4) photoproduct, an apurinic/apyrimidinic (AP) site or a 2-base pair mismatch [45]. This affinity is diminished by increasing the number of consecutive

base mismatches and essentially no binding occurs to single-stranded DNA [45], or to duplexes containing an 8-oxoguanine or O^6-methylguanine residue [46]. A common feature of the high-affinity substrates carrying (6-4) photoproducts or AP sites is the presence of an empty space between the lesion and the opposite bases [52]. This coincidence suggests that UV-DDB may discriminate between normal and damaged deoxyribonucleotides by penetrating into the double helix in search for an unusually wide spacing between complementary strands.

An early "hit-and-run" hypothesis suggested that DDB1 possesses damage-specific DNA-binding activity and that DDB2 leaves the complex after mediating the association of its larger partner with the DNA substrate [47]. Other studies indicated that UV-DDB activity is dependent on the presence of both subunits [45,53]. Yet different results have been obtained by Kulaksiz et al. [38], who tested the binding of each purified UV-DDB subunit to a 50-mer DNA duplex containing a centrally located (6-4) photoproduct. These authors came to the conclusion that the DNA-binding domain maps to the smaller DDB2 subunit. A separate study, performed in intact human fibroblasts, showed that the knockdown of DDB1 does not prevent DDB2 from accumulating at foci of UV damage, thus supporting the view that DDB2 is sufficient for target site binding [54]. A previous report already demonstrated that, upon UV irradiation, DDB1 translocates into the nucleus [55], but this nuclear accumulation is prevented by *DDB2* mutations in XP-E cells [56]. On the other hand, not only XP-E cells lacking DDB2 but also DDB1 knockdown cells are defective in photoproduct excision, indicating that both subunits are required for efficient repair [54].

The consequences of a DDB1 defociency may be explained by its role as an adaptor that connects a ubiquitin ligase complex to protein targets [57]. In general, the ubiquitylation reaction is directed by a three-enzyme cascade involving the ubiquitin-activating enzyme E1, a ubiquitin-conjugating enzyme (E2) and a ubiquitin protein ligase (E3). DDB1 forms a

molecular adaptor for the Cul4A-Roc1 ubiquitin ligase by mediating the recognition of WD40-repeat proteins, including DDB2 itself [58,59]. Known substrates of the DDB1-Cul4A-Roc1 machinery include DDB2 [60,61], XPC [62] and histones [63,64]. Ubiquitylated DDB2 looses its damaged DNA-binding activity [62] and is rapidly degraded [65], which contradicts its presumed role in the recognition of cyclobutane dimers because the majority of these lesions are still not repaired when most DDB2 is destroyed. Instead, the concurrent ubiquitylation of XPC is reversible and does not lead to protein degradation.

To summarize, DDB2 is thought to be the first factor that recognizes UV damage while its interaction partner, DDB1, mediates the physical handover of the lesions to the next recognition subunit. However, this model fails to explain the efficient removal of cyclobutane dimers in a number of *in vivo* or *in vitro* systems lacking DDB2 [38,66]. It may be possible that the ubiquitylation initiated by UV-DDB plays another role that is independent of NER activity. One intriguing observation is that skin fibroblasts taken from XP-E patients, or embryonic fibroblasts derived from $DDB2^{-/-}$ knockout mice, are more resistant than wild-type controls to UV-induced cell killing [67,68]. Further analysis of $DDB2^{-/-}$ cells revealed that this factor is involved in a regulatory circuit that controls the level of p53 in response to DNA damage. These results lend support to an alternative hypothesis whereby the primary role of UV-DDB is to trigger an apoptotic signaling cascade in response to genotoxic stress.

5. The special case of transcription-coupled repair

NER operates through two subpathways that differ in the initial recognition of base damage. In “global genome repair” (GGR), XPC, XPA, TFIIH and RPA bind to damaged sites and induce DNA incision irrespective of whether the target sequence is silent or actively engaged in transcription. However, living organisms have evolved a more efficient “transcription-coupled repair” (TCR) reaction that eliminates DNA lesions from the transcribed strand of

active genes [69]. Due to this specialized pathway, DNA excision repair is highly non-uniform in the chromosomal context. For example, cyclobutane dimers are removed more rapidly from transcribed genes than from transcriptionally silent regions [70,71], and DNA lesions are repaired in the template strand of RNA polymerase II-transcribed genes more rapidly than in the non-transcribed coding strand [72]. GGR and TCR play different biological roles: GGR protects from damage-induced mutations that ultimately lead to cancer, while TCR ensures that the genes are efficiently and correctly transcribed, thus protecting from premature aging [73].

The progression of RNA polymerase II along transcribed strands is obstructed by DNA modifications [see for example Ref. 74], indicating that, in TCR, the exquisite sensitivity of RNA polymerases to DNA lesions provides a facilitated mechanism of damage recognition. After immobilization at damaged sites, the stalled RNA polymerase II serves as a "bait" for the recruitment of XPA, TFIIH, RPA, as well as the endonucleases XPG and XPF [75], whereas UV-DDB and XPC are no longer necessary [76]. DNA repair only occurs if the RNA polymerase complex, which occludes the site of damage, is temporarily relieved. This poorly understood process involves the Cockayne syndrome (CS) complementation group A and B proteins [77]. The characteristic hallmark underlying the hereditary condition known as CS syndrome is a defect in the recovery of mRNA synthesis after UV irradiation [78] and this TCR deficit causes a clinical phenotype of postnatal growth failure, progressive neurodegeneration and symptoms reminiscent of segmental accelerated aging [79]. CSA is another WD40-repeat protein that acts as a cofactor for Cul4A-containing E3 ubiquitin ligases [59]. CSB, on the other hand, is a member of the SWI2/SNF2 family of DNA-dependent ATPases with chromatin remodeling activity [80,81].

Several hypotheses have been raised regarding the TCR mechanism. For example, the arrested RNA polymerase II is phosphorylated and subsequently polyubiquitylated by a

reaction that involves the CSA and CSB proteins [82,83] and two interpretations of the role of ubiquitylation have been proposed. One indicates that ubiquitin marks the RNA polymerase II molecule for degradation, leaving the damaged strand accessible for repair [84]. The other hypothesis suggests that a blocked RNA polymerase II molecule does not need to be degraded and that ubiquitylation is a signal for activation of DNA repair or other cellular responses [85]. More recent findings indicate that CSB, TFIIH and XPG cooperate to remodel the RNA polymerase II complex in an ATP-dependent manner. The resulting conformational shift would allow access to the lesion without removal or disruption of the transcription machinery [86].

6. Bipartite substrate discrimination in the GGR pathway

How does the GGR system, without any help from the transcriptional machinery, detect DNA damage in a versatile manner and, at the same time, avoid futile repair cycles among the 3 billion base pairs of the human genome? A major decision point during the sequential assembly of NER complexes is related to the question of how the system "knows" whether cleavage should occur, which is appropriate only if a lesion is actually present. This fundamental problem of versatile DNA damage recognition is solved by a bipartite discrimination strategy that employs several distinct subunits to detect different characteristic features of damaged DNA.

Hanawalt and Haynes [87] were the first to propose that the need for excision repair is determined by comparing the conformation of damaged DNA to that of the normal Watson-Crick double helix. Elaborating on this concept, Gunz et al. [88] showed that the efficiency of bulky lesion recognition by the human GGR complex correlates with the degree of helical destabilization arising from the loss of base pairing properties at DNA adducts. It was, therefore, expected that the GGR factors responsible for the initial damage recognition step

would show an affinity for helical distortions caused by DNA lesions. Simple base mismatches or nucleotide bulges are, however, not processed by the GGR machinery, indicating that the local thermodynamic destabilization of duplex DNA is not sufficient to qualify as a NER substrate [89,90].

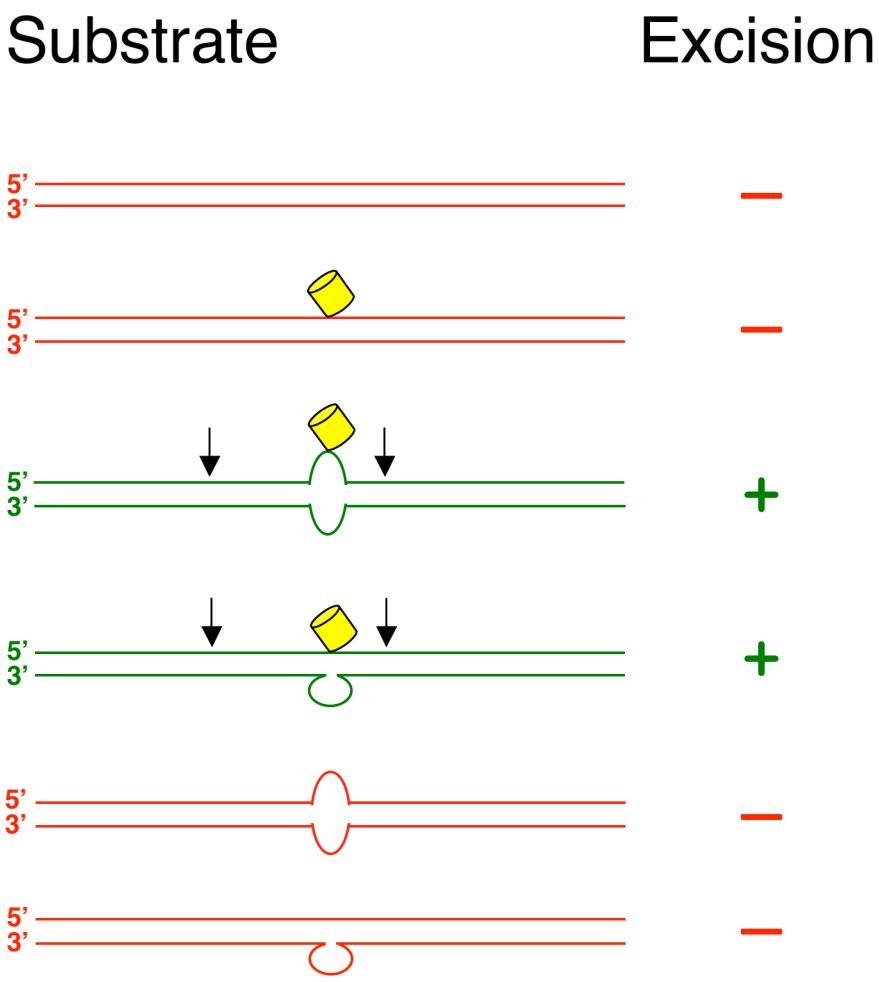


Figure 2. Substrates designed to demonstrate the concept of bipartite substrate discrimination in the human NER pathway. An artificial "non-distorting" adduct is recognized by the NER system only in combination with a concomitant distortion induced by 1-3 base mismatches or a short DNA loop. However, undamaged homoduplexes or heteroduplexes containing

mismatches or loops do not constitute a substrate for the NER process. The sites of incision are indicated by the arrows.

Understanding the strategy used by the NER system to discriminate damaged sites required the construction of highly defined substrates amenable to molecular manipulations. In fact, the notion of bipartite substrate discrimination originated from *in vitro* excision assays demonstrating that the human GGR complex remains inactive on DNA duplexes containing a "non-distorting" DNA adduct that preserves normal hydrogen bonds between complementary bases [91]. As indicated in Fig. 2, however, such a "non-distorting" DNA adduct in conjunction with local disruption of canonical base pair interactions, caused by mismatches or a DNA bulge, induces strong NER reactions. Thus, these experiments revealed that the molecular hallmark leading to GGR activity consists of two distinct elements, i.e., disruption of Watson-Crick base pairing and altered chemistry of the damaged deoxyribonucleotide residue [91,92]. Neither defective base pairing alone, in the absence of bulky adducts, nor defective chemistry in the absence of helical distortions, is able to elicit an excision response, but the combination of these two substrate alterations results in the assembly of productive excision complexes. The term of "bipartite recognition" has been introduced to indicate that the GGR factors use at least two principal levels of discrimination to recognize damaged substrates.

Further support for a bipartite mechanism of substrate discrimination was sought by studying the excision of DNA adducts caused by bay-region benzo[*a*]pyrene (B[*a*]P) and the related fjord-region benzo[*c*]phenanthrene (B[*c*]Ph) diol epoxides [93]. A fjord-region (+)-*trans*-*anti*-B[*c*]Ph-N^6-dA adduct, which retains normal Watson-Crick base pairing, is not excised during incubation in human cell extracts when situated in a fully complementary duplex. However, the same (+)-*trans*-*anti*-B[*c*]Ph-N^6-dA adduct in combination with three

mismatched bases stimulates DNA excision in human cell extracts (Fig. 3). Such a compound substrate, containing a fjord adduct and base mismatches, is excised with kinetics similar to the helix-destabilizing (+)-*trans*-*anti*-B[*a*]P-N^6-dA bay adduct [94]. The markedly more efficient repair of *cis*-*anti*-B[*a*]P-N^2-dG adducts, as compared to excision of the stereoisomeric *trans*-*anti*-B[*a*]P-N^2-dG adducts, is also in accord with the hypothesis of bipartite recognition, since the Watson-Crick geometry between base pairs is disrupted in the *cis* conformation but not in the *trans* conformation [95,96].

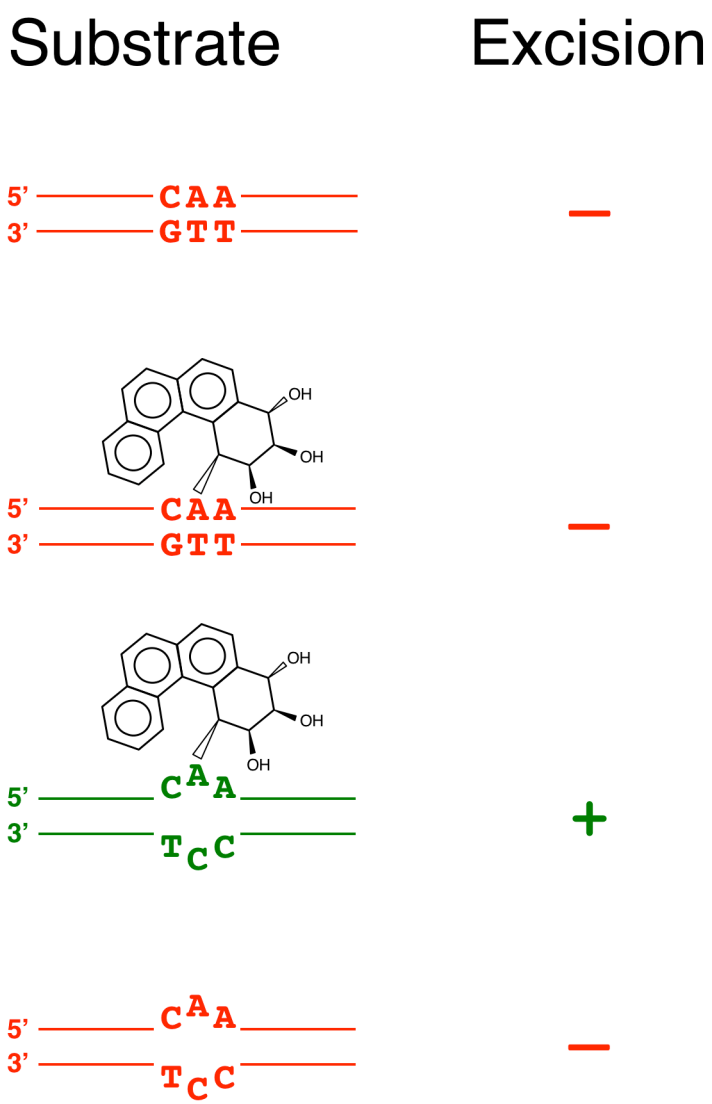


Figure 3. Demonstration of bipartite recognition scheme in the presence of DNA lesion induced by a known carcinogen. The non-distorting benzo[*c*]phenanthrene adduct is recognized only in combination with base mismatches. However, a bulge generated by base mismatches, in the absence of DNA adducts, is not a NER substrate.

7. Common conformational features of NER substrates

As mentioned before, the NER system is able to process diverse UV lesions, including the more abundant cyclobutane pyrimidine dimers and the less frequent (6-4) photoproducts, as well as wide range of bulky carcinogen-DNA adducts, oxidative lesions, crosslinked purines, protein-DNA crosslinks and other modifications that share no overt structural similarity [88,97]. To reach this striking substrate versatility, the NER proteins are thought to recognize a conformational distortion of the double helix induced by UV photoproducts and other types of DNA damage [98]. An increased flexibility of damaged DNA duplexes, relative to the undamaged double helix, may provide such a generic property of different NER substrates. In fact, base stacking is the predominant energetic force leading to the intrinsic rigidity of DNA [99] but the loss of base stacking, resulting in a flexible hinge, is a common consequence of bulky lesion formation [100,101].

Even the native double-stranded DNA is not a static molecule and the DNA strands are constantly in motion due to thermal oscillations, such that the distance between complementary strands exhibits fast and small variations [102-104]. In the absence of DNA damage, the picosecond-to-nanosecond timescale of these small strand vibrations is probably too short to be recognized by DNA repair factors. However, molecular dynamics simulations predict that the introduction of a single lesion, for example a cyclobutane pyrimidine dimer, provokes longer-lived and larger openings of the double helix relative to undamaged DNA [105]. In the case of pyrimidine dimers, the covalently bonded residues move together in

phase, forcing the undamaged bases in the opposite strand to synchronize and give rise to more prominent oscillations compared to native sites. These large fluctuations between complementary strands appear 25 times more frequently at a cyclobutane dimer position than in undamaged DNA sequences. Also, the amplitude of these oscillations is drastically increased because the strength of interactions between the two complementary strands is weakened. Interestingly, these dynamic changes triggered by base damage generate mainly oscillations of the intact complementary sequence across bulky lesions, as the strand containing base adducts is less flexible than native DNA [105]. Thus, the simulation of macromolecular dynamics lends support to the hypothesis that damage-induced DNA fluctuations may provide a truly universal signal for the recruitment of repair factors.

8. XPC protein is a sensor of abnormal strand oscillations

It has been suggested that XPC protein, one of the key initiators of the GGR pathway, is sensitive to damage-induced deformations of the double helix [31,34], but the molecular basis of this recognition function remained unknown for a long time. The XPC subunit (125 kDa) is found in complexes with HR23B, a 58-kDa homolog of the yeast NER protein RAD23 [106], and centrin-2 (CTN2), a 18-kDa calcium-binding protein [16]. XPC protein possesses DNA-binding activity, whereas the HR23B and CTN2 partners exert accessory functions in stabilizing the complex and stimulating its action in DNA repair [34,107]. XPC protein alone or in conjunction with HR23B binds preferentially to damaged DNA substrates containing, for example, (6-4) photoproducts, acetylaminofluorene adducts or cisplatin crosslinks [31,34,108,109]. Scanning force microscopy studies showed that the binding of XPC protein to damaged double-stranded DNA induces a kink in the nucleic acid backbone [110]. The structural determinants for the recruitment XPC protein have been probed with artificial substrates, thus revealing a general affinity for sites that deviate from the canonical

Watson-Crick geometry, including single-stranded loops, mismatched bubbles or single-stranded overhangs [34,90,111].

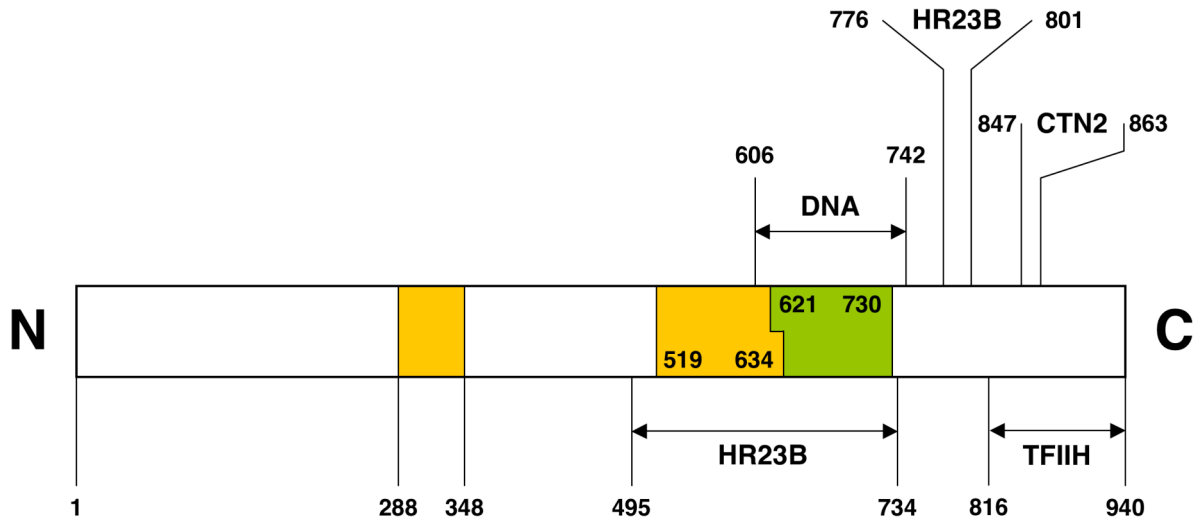


Fig. 4 Domain structure of human XPC protein. Yellow, homology with the family of peptide-*N*-glycanases; green, homology with OB-fold motifs. A minimal DNA-binding fragment has been reported in Ref. 113. Domains for the interaction with HR23B have been identified in Refs. 112 –114. The interactions with CTN2 and TFIIH are described in Refs. 16 and 113.

There is ambiguity over the precise amino acid region of XPC protein involved in the complex formation with HR23B. A two-hybrid study reported by Li et al. [112] mapped the HR23B-interacting region of XPC to residues 776 through 801 (Fig. 4). In contrast, Uchida et al. [113] performed a bidirectional truncation study to map the minimal HR23B-interacting region of XPC between amino acids 496 and 734. A more recent report using a series of XPC fragments expressed in bacteria proposed that there is an additional HR23B-binding site in the N-terminal region of XPC protein [114]. This N-terminal domain is also responsible for an interaction with XPA [114], which may mediate the transition from an initial recognition

intermediate (involving XPC and TFIIH) to the formation of an ultimate incision complex that includes the two endonucleases [115]. In fact, XPC protein behaves like a "molecular matchmaker" as it initiates the assembly of a repair complex but leaves the DNA substrate before completion of the incision reaction [15,116]. The carboxy-terminal tail of XPC protein mediates the association with CTN2 (residues 847-863) and with TFIIH (residues 816-940) [16,113].

Analysis of the XPC sequence revealed an intriguing similarity to the transglutaminase fold of peptide-*N*-glycanases, which remove glycan modifications from glycoproteins during their degradation. However, the amino acid sequence of XPC protein lacks the predicted catalytic triad (Cys-His-Asp) characteristic of this family of enzymes, suggesting that this DNA repair subunit emerged during eukaryotic evolution through duplication of an ancient peptide-*N*-glycanase, followed by the loss of enzymatic activity [117]. The regions of homology with peptide-*N*-glycanases include residues 288-348 and 519-634 of the human XPC protein and these domains (Fig. 4) are thought to be involved in protein-protein interactions. On the other hand, we recently discovered a sequence homology between RPA-B, one of the single-stranded DNA-binding domains of human RPA, and an XPC region extending from residue 621 to 730 [118]. The observed homology (27% amino acid identity and 73% similarity) includes most of the conserved elements of secondary structure characteristic of the oligonucleotide/oligosaccharide-binding fold (OB-fold) responsible for the tight interaction of RPA with single-stranded DNA. A sequence similarity of 64% and 66%, respectively, has also been observed between XPC residues 621-730 and two distinct OB-folds of BRCA2.

The single-stranded DNA-binding activity of each OB-fold in RPA or BRCA2 correlates with the presence of two structurally conserved aromatic side chains that mediate stacking interactions with closely spaced DNA bases [119]. The search for functionally analogous

aromatics in human XPC protein revealed that Trp690 and Phe733 are critically required for DNA binding and GGR activity [118]. Consistent with the presence of a putative OB-fold motif, XPC protein displays a preference for single-stranded oligonucleotides, implying that it recognizes the local single-stranded character of DNA containing bulky lesions. Surprisingly, XPC exhibits an unfavorable binding to damaged oligonucleotides compared to the more efficient interaction with undamaged single-stranded counterparts [109,118]. This finding is consistent with an indirect mode of bulky lesion recognition that exploits the local loss of normal duplex properties, resulting in the appearance of a single-stranded character in the undamaged complementary sequence opposite to the damage.

To summarize, the striking affinity of XPC protein for single-stranded oligonucleotides, in combination with its aversion to interact with damaged single strands, indicates that one of the early recognition steps in the GGR pathway is guided by an association with the native strand of damaged duplexes. This mechanism of action fits with the appearance of large and long-lived oscillation in the native DNA strand across lesion sites, thus predicting that XPC protein operates at sites of bulky lesions by capturing the local and transient formation of a single-stranded conformation in the undamaged complementary sequence (Fig. 5). One advantage of this inverted model of substrate discrimination is that the early recognition step is independent of the variable chemistry of lesion sites and, hence, contributes to the ability of the GGR machinery to detect a very wide array of DNA adducts. The affinity of XPC protein for abnormally oscillating DNA sequences may also be used to facilitate other excision repair processes, as XPC has been shown to interact with 3-methyladenine DNA glycosylase [120], thymine DNA glycosylase [121] and 8-oxoguanine DNA glycosylase [10]. Accordingly, the XPC-HR23B-CTN2 complex may constitute a platform not only for the loading of GGR factors onto damaged DNA, but also for the recruitment of a battery of enzymes involved in base excision pathways.

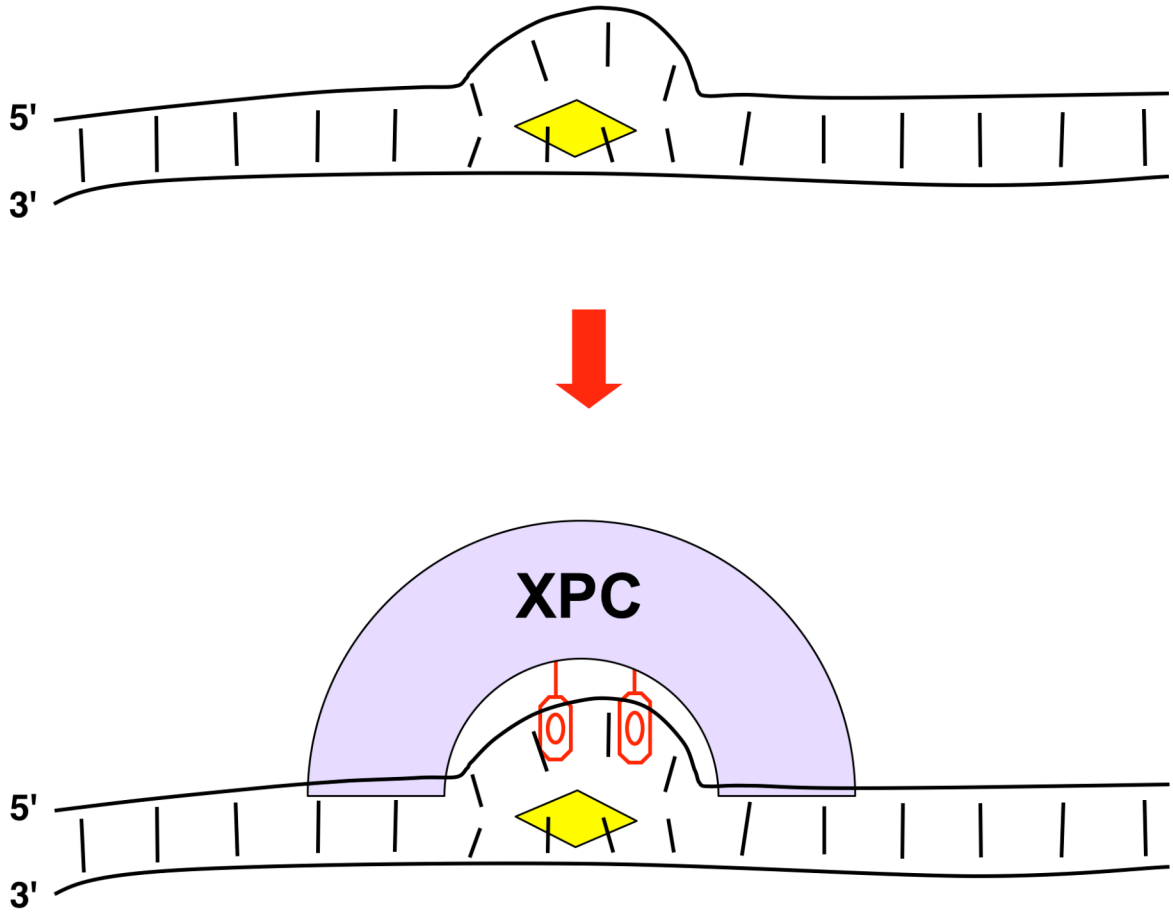


Fig. 5 Scheme illustrating how human XPC protein may deploy aromatic side chains (shown in red color) to capture DNA oscillations in the undamaged strand opposite to bulky lesions. According to this model, the unpaired residues opposite to the lesion are sandwiched between aromatic side chains of XPC protein.

9. XPA protein is a sensor of helix bendability

When migrating in denaturing gels, XPA protein forms several bands with an apparent mass of 40-45 kDa. The discrepancy with the calculated molecular mass (31 kDa) has been ascribed to the presence of disordered regions in the polypeptide fold. Also, the presence of multiple electrophoretic bands is thought to reflect distinct polypeptide conformations and

this molecular flexibility of XPA has been related to its function in accommodating a disparate variety of lesions [122,123]. In fact, internal motions may alter the nucleic acid interaction surface to fit different kinds of damaged DNA substrates [124]. The retention time of recombinant human XPA protein in gel filtration experiments suggests the formation of homodimers in solution [125], but *in vivo* studies indicate that the majority of XPA molecules diffuse rapidly in monomeric form within the nuclear compartment [126].

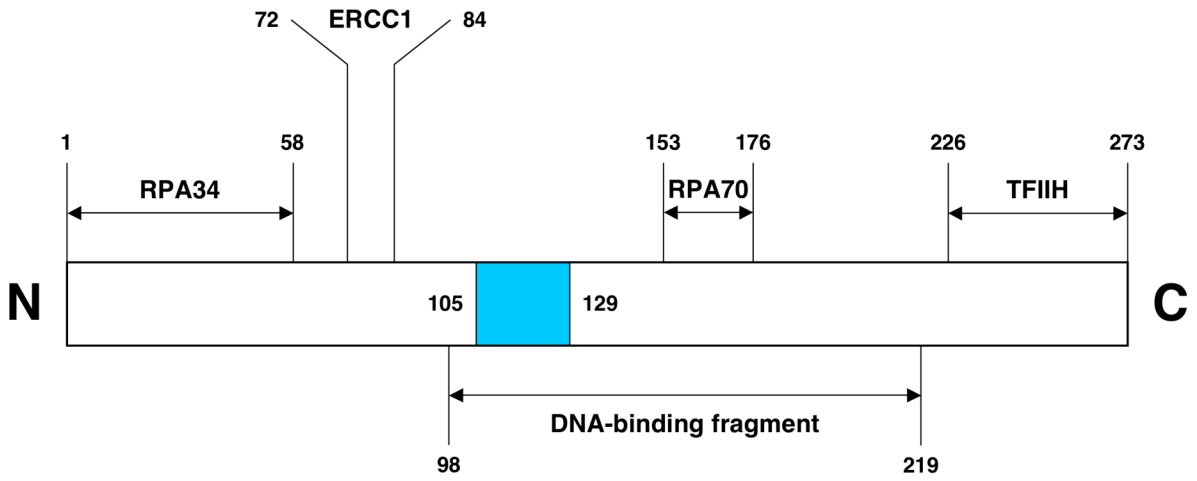


Fig. 6 Domain structure of human XPA protein. Blue, zinc finger domain. The minimal DNA-binding fragment is described in Refs. 141-143. Domains for the interaction with the 34- and 70-kDa subunits of RPA, with ERCC1 and TFIIH have been identified in Refs. 127-131.

XPA associates with several other NER subunits and specific interaction domains have been identified by deletions studies (Fig. 6). The N-terminal portion (residues 1-58) and a central region (residues 153-176) contain sequences for binding to RPA. An association of XPA with the large 70-kDa subunit of RPA (RPA70) is essential for the NER function [127]. The C-terminal region (residues 226-273) binds to TFIIH [128]. XPA also forms complexes

with the ERCC1-XPF heterodimer, and amino acid sequences involved in the interaction with this endonuclease are the polyglutamic acid cluster (residues 78-84) as well as a nearby tetrapeptide consisting of residues 72-75 [129-131].

The DNA-binding activity of XPA is characterized by some selectivity for UV- or chemical carcinogen-damaged duplexes [132]. Compared to UV-DDB and XPC, however, the affinity of XPA for damaged duplexes is orders of magnitude lower. For example, Jones and Wood [133] estimated the binding constant of XPA for (6-4) photoproducts in double-stranded DNA to be ~$3 \cdot 10^6$ M, whereas the reaction constant for binding of UV-DDB to the same substrate is > 5'000-fold greater. Attempts to produce footprints of XPA protein on damaged DNA by nuclease protection or other techniques have failed. Nevertheless, a DNA damage "verification" function has been proposed for XPA because its affinity for nucleic acid substrates is increased in conjunction with RPA [127,134-136] or ERCC1 [137]. These larger complexes bind to damaged DNA more avidly than each protein alone, indicating that the different subunits may cooperate to promote substrate recognition. XPA also interacts with XPC protein, but it is not yet clear whether this association stabilizes the recognition complex [33] or results in the displacement of XPC from damaged DNA [115].

XPA protein also displays an affinity for distorted DNA structures carrying mismatches, loops or bubbles, even if no actual DNA lesion is introduced into the substrate [135]. It has a particularly strong preference for distorted DNA molecules, such as three- or four-way DNA junctions, which share the architectural feature of presenting two double strands emerging from a central bend [135,138]. In view of these findings, we propose that XPA serves to recognize a DNA kink resulting from the inherently increased DNA flexibility at lesion sites (Fig. 7). Because sharp DNA bends are often introduced when multi-protein machines assemble on DNA [139], a local kink is likely to be further stabilized in the context of NER

complexes [139]. The notion that a site-specific kink may be formed during bulky lesion recognition is further supported by the paradigm of NER in prokaryotic organisms [140].

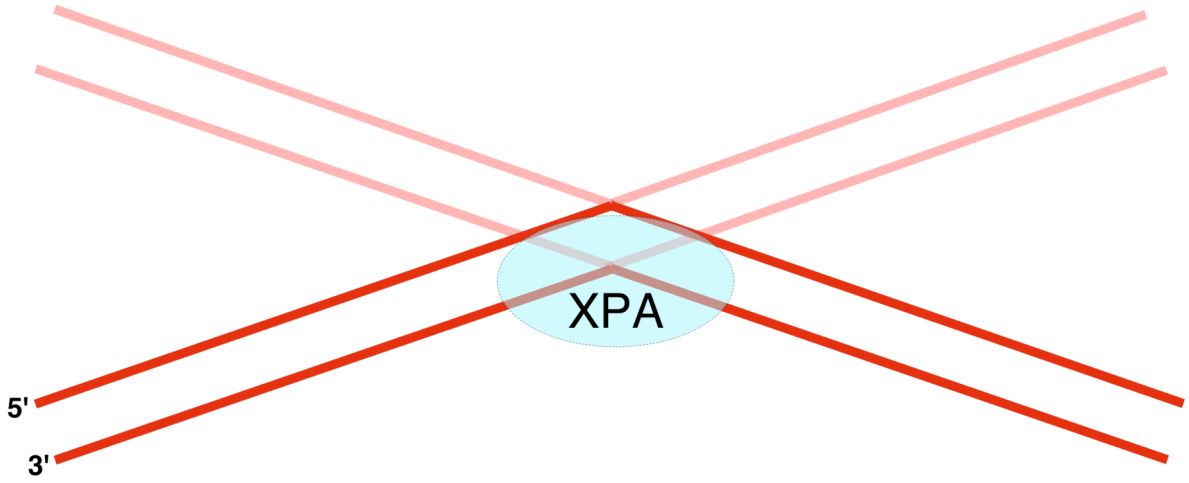


Fig. 7. Recognition of DNA kinks by XPA protein. Four-way junctions have been used as model substrates to mimic sharply bent helical backbones.

A nucleic acid interaction domain has been identified by nuclear magnetic resonance (NMR) spectroscopy [141,142]. This solution structure analysis revealed that the central region of 122 amino acids (residues 98-219) is composed of an acidic subdomain (residues 105-129) containing a zinc finger, and a C-terminal subdomain (residues 138-209) that forms a positively charged cleft on the protein surface. Subsequent chemical shift perturbation experiments conducted in the presence of either DNA or a short RPA peptide sequence led to the unexpected finding that the zinc finger domain is not involved in DNA binding but, instead, is required for the interaction with RPA. Conversely, the cationic cleft has the appropriate curvature and size to accommodate DNA [141,143], thus prompting a mutational screen to determine the functional role of each basic residues in this presumed DNA-binding site. The characterization of mutants by electrophoretic mobility shift, photocrosslinking and

host-cell reactivation assays demonstrated that positively charged side chains on the XPA surface are indeed required for the efficient interaction with target DNA. In particular, two neighboring basic residues (Lys179 and Lys141), on the N-terminal side of the DNA-binding cleft, form a dual hotspot for recognition of the nucleic acid substrate [144].

Critical lysines have been converted to negative moieties by mutating these amino acids to glutamic acid. Interestingly, the Lys179Glu/Lys141Glu tandem substitution conferred a stronger DNA repair defect than any other combination of double mutants throughout the DNA-binding surface [138]. Like the respective single mutants, this Lys179Glu/Lys141Glu double mutant fails to interact with linear DNA fragments but is still able to bind to 3- or 4-way DNA junction molecules used as a surrogate for kinked substrates. Surprisingly, the Lys179Glu/Lys141Glu tandem mutant binds to four-way DNA junctions with exactly the same affinity as wild-type XPA, although it generates nucleoprotein products that migrate faster in native gels than the control complexes generated by wild-type protein. Photocrosslinking experiments revealed that the subtle molecular defect underlying the formation of such abnormal complexes resides in the inability of the Lys179Glu/Lys141Glu double mutant to undergo close contacts with the kinked junction region of the tested DNA structures. Also, unlike wild-type XPA, the aberrant nucleoprotein complexes formed by the tandem mutant are unable to recruit the XPF-ERCC1 endonuclease [138]. In combination, these results lend support to the hypothesis that the assembly of a productive incision intermediate, which includes XPF-ERCC1, is dependent on the proper association of XPA protein with the bending angle induced by site-specific kinks in the DNA substrate.

To conclude, the requirement for positively charged side chains in the DNA-binding surface, and the preference for kinked DNA, indicates that XPA may represent a molecular sensor of abnormal electrostatic potentials in the nucleic acid substrate. The case of UV endonuclease V, for which detailed crystallographic data is available, illustrates that a cluster

of basic amino acids may participate in damage recognition through electrostatic interactions with the unique backbone deformation induced by a DNA kink [145]. This sensor mechanism detects the higher density of negative charges arising from the closer spacing of phosphate moieties at narrow DNA bends.

10. TFIIH is a sensor of defective deoxyribonucleotide chemistry

The TFIIH complex shuttles between sites of transcription by RNA polymerase I or II, and sites of excision repair [27]. TFIIH can be resolved in two main components: the core complex consisting of 6 polypeptides (XPB, TTDA, p62, p52, p44, p34), which assemble in a ring-like structure with a central hole [146,147] and a protruding CAK (Cdk-activating kinases) complex containing cdk7, cyclin H and MAT1 [148]. This CAK component is not required for the NER function. XPD protein, which appears to play a crucial role in damage recognition, is found in both subcomplexes. In transcription, DNA unwinding by TFIIH allows the nascent RNA molecules to escape from the promoter region and progress towards the elongation phase [149]. In the GGR process, TFIIH is presumably recruited to GGR sites by XPC protein through interactions with the XPB and p62 subunits [150]. TFIIH then separates the two strands around the lesion, until an approximately 30-nucleotide "bubble" is formed. This unwinding activity generates an open intermediate characterized by double-stranded to single-stranded transitions on either side of the lesion, thus providing the substrates for double DNA incision by structure-specific endonucleases [17,20].

Central to the local unwinding process are the two DNA helicases, XPD and XPB. The XPD subunit functions primarily in DNA repair because its helicase activity, which has 5' to 3' polarity, is required for NER but is dispensable for transcription [151]. It is thought that XPD has a more structural role in transcription by acting as a bridge between the core TFIIH ring and the CAK protrusion [146]. XPB unwinds double-stranded DNA with opposite 3' to

5’ polarity and this activity is required for both transcription and the NER process [152]. By analogy with the prokaryotic recombination complex RecBCD [153], a bipolar pair of DNA helicases may serve to unwind the double helix by a joint mechanism whereby the two enzymes translocate with opposite polarity, but in the same direction, on each strand of the antiparallel DNA duplex (Fig. 8A). When tested in bidirectional DNA unwinding assays using undamaged substrates, the whole TFIIH complex containing both XPD and XPB only displays 5’ to 3’ polarity and a side-by-side comparison between purified recombinant proteins showed that XPD has a much higher specific activity than XPB [154]. These results converge on the hypothesis that the XPD subunit provides the main molecular engine that drives translocation of the entire TFIIH complex along DNA (Fig. 8A).

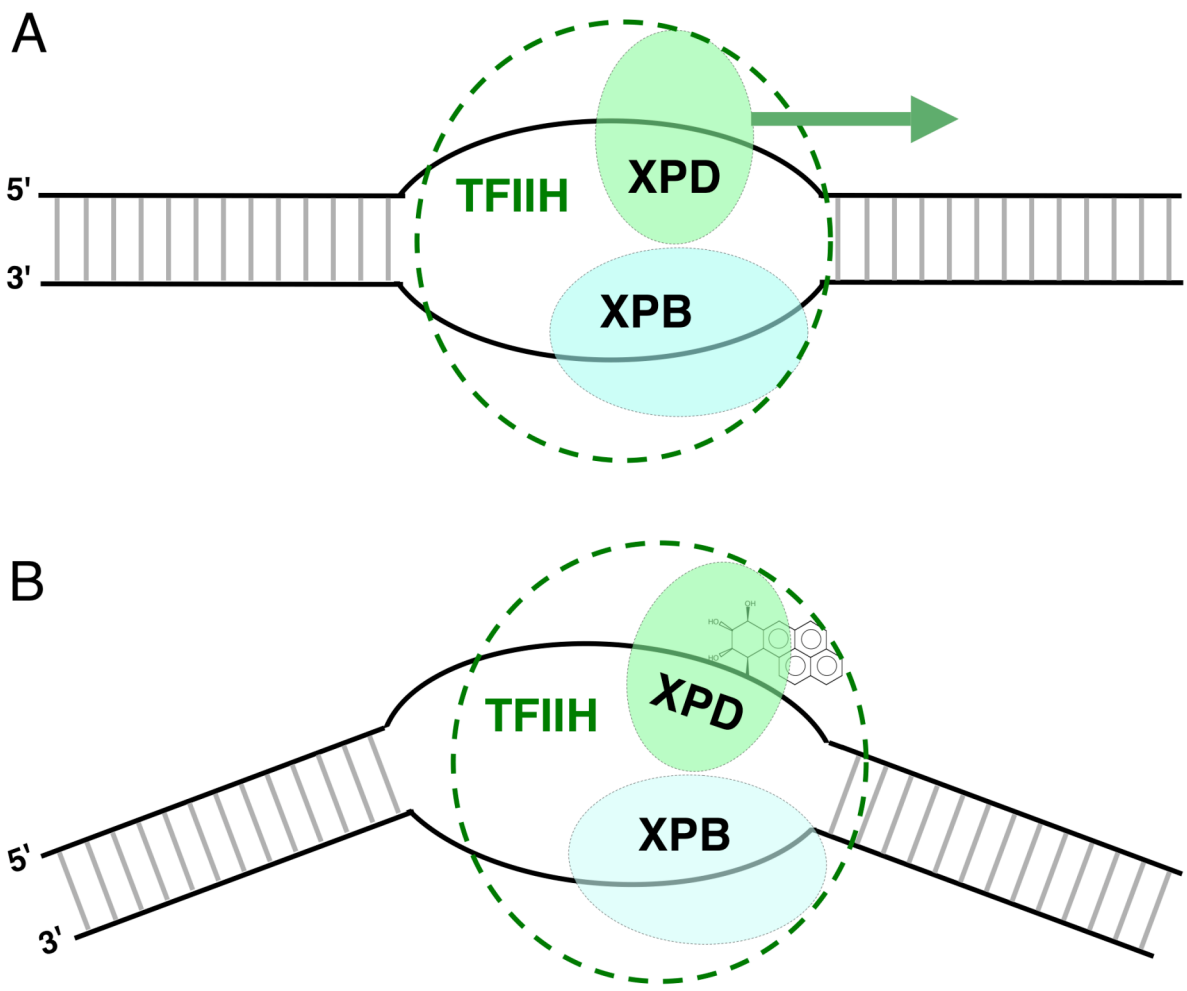


Fig. 8 Model depicting how two DNA helicases with opposite polarity may cooperate to unwind double-stranded DNA and locate defective deoxyribonucleotide residues.

(A) Action of DNA helicases on undamaged duplex substrates.

(B) Response to damaged substrate containing an obstructing DNA adduct. Inhibition of XPD helicase generates an open intermediate with double-straned to single-stranded junctions flanking the lesion.

Since the bases are buried inside the double helix, the two DNA strands have to be separated for the efficient localization of base lesions. The initial interaction of XPC protein with the undamaged strand may facilitate the subsequent loading of XPD onto the damaged strand, such that this DNA helicase constitutes the first subunit that comes in direct contact with the offending residue. Such a role of XPD in the detection of DNA damage is suggested by an analogy with the bacterial DNA helicase UvrB, a recognition subunit of the prokaryotic NER pathway. Whereas XPD and UvrB do not share an overall sequence identity, both proteins display conserved helicase motifs that provide a common scaffold for structural comparisons [155]. Thus, a model for the human XPD protein has been established based on its similarities to the bacterial UvrB subunit, for which high-resolution crystal structures are available [156-158]. This model suggests that XPD, like UvrB, deploys a β-hairpin domain that, once inserted between the strands of the duplex, is able to sense the presence of DNA lesions [155,159].

Using a yeast ortholog of XPD, the Rad3 helicase of *Saccharomyces cerevisiae*, we have previously demonstrated that this molecular engine is arrested by DNA lesions located on the strand along which the enzyme is translocating. Like XPD, purified Rad3 protein requires single-stranded regions to initiate unwinding and this loading strand must be free of damage to promote full ATPase and DNA helicase activity. If, however, the loading strand contains

DNA lesions, both enzymatic activities cease when Rad3 protein encounters the chemically altered residue. Furthermore, the presence of base damage induces the formation of stable Rad3 protein-DNA complexes, indicating that Rad3 protein becomes sequestered on DNA at lesion sites [160,161]. In contrast, lesions in the opposite complementary strand have no effect on this tracking mechanism [160].

As described before, the hypothesis that XPD and the yeast homolog Rad3 may participate directly in the detection of DNA damage is supported by a site-directed crosslinking study revealing that XPD is located in close proximity to bulky lesions within the ultimate excision complex [35]. It seems intuitive to propose that inhibition of the 5' to 3' helicase activity serves to localize damaged deoxyribonucleotides during the NER process. Fig. 8B illustrates how the strand-specific block of one of the two DNA helicases (XPD) may produce a stable recognition complex in which the DNA becomes denatured and contorted around the lesion site. This enzymatic scanning function provides a mechanism to confirm the presence of bulky lesions before recruiting the incision endonucleases.

11. Role of XPA-RPA interactions in the open recognition complex

XPA is not necessary for the damage-specific recruitment of RPA [126], indicating that the majority of XPA and RPA molecules are not interacting in the absence of DNA damage and that a co-localization of these two subunits only occurs at their site of action in the ultimate excision complex. The independent recruitment of XPA and RPA to lesion sites raises the question of why, as reported by Li et al. [127], the interaction between these two partners is essential for the NER reaction. RPA represents the most abundant single-stranded DNA-binding factor in human cells. Each RPA monomer occupies approximately 30 nucleotides [162], which corresponds roughly to the length of the gapped DNA intermediate generated during the NER process. This coincidence between the size of excision products and the

occupancy length on single-stranded DNA implies that RPA might protect the undamaged intact strand from inadvertent nuclease attack [163,164]. In fact, RPA interacts preferentially with the undamaged strand and this bias is further increased by the addition of its interaction partner XPA [165,166].

We have previously proposed that XPA-RPA interactions are required to further verify the need for endonucleolytic processing of the DNA substrate [135]. According to this model, XPA and RPA would adopt distinct roles to "double-check" the NER pathway: a lesion verification function is attributable to XPA, whereas RPA is recruited to ensure that the two DNA strands are separated from each other and to protect the undamaged template. Once RPA is bound to single-stranded DNA opposite the lesion, it provides a docking platform together with XPA to position the structure-specific endonucleases XPF-ERCC1 (which makes the 5' incision) and XPG (which makes the 3' incision) in a proper orientation [22-24]. Thus, both endonuclease subunits are dependent on XPA-RPA interactions for their correct three-dimensional localization to the respective sites of action in the vicinity of DNA adducts.

12. Conclusion: bipartite DNA damage recognition in the GGR pathway

Lin and Sancar [167] predicted several years ago that damage recognition in NER processes is achieved by a selectivity cascade, in which different steps of low selectivity in succession lead to a specificity that is comparable to that of transcription factors or other sequence-specific DNA-binding proteins. Based on the evidence that has been reviewed here, we elaborate on this hypothesis to propose a bipartite substrate discrimination mechanism, in which UV-DDB, XPC and XPA constitute the sensors of abnormal DNA conformations, whereas TFIIH functions as a tracking enzyme that locates the chemically damaged residues.

A salient feature of this bipartite model of substrate discrimination is that the early sensors of DNA damage avoid direct contacts with bulky lesions and, instead, recognize deformations of the double helix. Numerous studies have shown that the efficiency of bulky lesion excision depends on the extent of base pair destabilization in the immediate vicinity to the damaged nucleotide [34, 88, 90, 91, 93]. Dynamic simulations reveal that a characteristic property of damaged DNA is the presence of abnormal oscillations between the complementary strands of the double helix [105]. An increased bendability has been identified as another common property of damaged substrates containing bulky lesions [100]. Thus, DNA damage recognition begins when UV-DDB, XPC and XPA probe the thermodynamic stability of the double helix and detect abnormal dynamic fluctuations.

XPC protein attracts TFIIH and loads the ring-like helicase domain of TFIIH onto the damaged strand. Driven by the 5' to 3' helicase activity of XPD protein, TFIIH moves up to a distance of about 15 nucleotides along the damaged strand in search of the lesion [92]. The activity of TFIIH promotes partial unwinding by 20-25 base pairs, thereby separating the duplex. This tracking activity serves to probe the chemical composition of the target strand and to determine the precise location of the adducted nucleotides. Damage recognition is completed when XPD encounters the adduct and becomes sequestered on the damaged strand. The TFIIH complex intervenes in the reaction in a way that erroneous intermediates can be aborted before generating spurious incisions. If, however, the assembly occurs accidentally at an undamaged site, ATP hydrolysis by TFIIH leads to dissociation of the repair intermediate. An similar proofreading activity has been demonstrated for the mismatch repair factor MutS. In the presence of homoduplex DNA, MutS hydrolyzes ATP and dissociates from the nucleic acid substrate. In the presence of a base mismatch, ATP hydrolysis is inhibited, which allows the stable MutS-DNA complex to form [168]. By analogy, the unhindered burst of ATP hydrolysis by XPD may lead to dissociation of the

initial recognition complex from undamaged DNA. However, progression of TFIIH is blocked by bulky lesions, thereby marking the location of the genetic insult. To conclude, this bipartite mechanism for lesion recognition not only results in the ability to detect a wide range of different DNA lesions but, at the same time, also protects undamaged DNA, including the complementary template strand across lesion sites, from inadvertent incisions.

Acknowledgements

H.N. is supported by the Swiss National Science Foundation (grant 3100A0-113694). The work of K.B.B. is partly supported by the Center for Nonlinear Studies at the Los Alamos National Laboratory.

References

[1] E.C. Friedberg, G.C. Walker, W. Siede, R.D. Wood, R.A. Schultz, T. Ellenberger, DNA Repair and Mutagenesis, ASM Press, Washington D.C., 2006

[2] A. Sancar, DNA excision repair, Annu. Rev. Biochem. 65 (1996) 43-81

[3] R.D. Wood, Nucleotide excision repair in mammalian cells, J. Biol. Chem. 272 (1997) 23465-23468

[4] W.L. de Laat, N.G. Jaspers, J.H. Hoeijmakers, Molecular mechanism of nucleotide excision repair, Genes Dev. 13 (1999) 768-785

[5] M.S. Satoh, C.J. Jones, R.D. Wood, T. Lindahl, DNA excision-repair defect of xeroderma pigmentosum prevents removal of a class of oxygen free radical-induced base lesions, Proc. Natl. Acad. Sci. USA 90 (1993) 6335-6339

[6] J.T. Reardon, T. Bessho, H.C. Kung, P.H. Bolton, A. Sancar, In vitro repair of oxidative DNA damage by human nucleotide excision repair system: possible explanation for neurodegeneration in xeroderma pigmentosum patients, Proc. Natl. Acad. Sci. USA 94 (1997) 9463-9468

[7] I. Kuraoka, C. Bender, A. Romieu, J. Cadet, R.D. Wood, T. Lindahl, Removal of oxygen free-radical-induced 5',8-purine cyclodeoxynucleosides from DNA by the nucleotide excision-repair pathway in human cells, Proc. Natl. Acad. Sci. USA 97 (2000) 3832-3837

[8] J.T. Reardon, A. Sancar, Repair of DNA-polypeptide crosslinks by human excision nuclease, Proc. Natl. Acad. Sci. USA 103 (2006) 4056-4061

[9] K.H. Kraemer, M.M. Lee, J. Scotto, DNA repair protects against cutaneous and internal neoplasia: evidence from xeroderma pigmentosum, Carcinogenesis 5 (1984) 511-514

[10] M. D'Errico, E. Parlanti, M. Teson, B.M. de Jesus, P. Degan, A. Calcagnile, P. Jaruga, M. Bjoras, M. Crescenzi, A.M. Pedrini, J.M. Egly, G. Zambruno, M. Stefanini, M. Dizdaroglu, E. Dogliotti, New functions of XPC in the protection of human skin cells from oxidative damage, EMBO J. 25 (2006) 4305-4315

[11] A. Aboussekhra, M. Biggerstaff, M.K. Shivji, J.A. Vilpo, V. Moncollin, V.N. Podust, M. Protic, U. Hubscher, J.M. Egly, R.D. Wood, Mammalian DNA nucleotide excision repair reconstituted with purified protein components, Cell 80 (1995) 859-868

[12] D. Mu, C.H. Park, T. Matsunaga, D.S. Hsu, J.T. Reardon, A. Sancar, Reconstitution of human DNA repair excision nuclease in a highly defined system, J. Biol. Chem. 270 (1995) 2415-2418

[13] S.J. Araujo, F. Tirode, F. Coin, H. Pospiech, J.E. Syvaoja , M. Stucki, U. Hubscher, J.M. Egly, R.D. Wood, Nucleotide excision repair of DNA with recombinant human

proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK, Genes Dev. 14 (2000) 349-359

[14] M. Volker, M.J. Mone, P. Karmakar, A. van Hoffen, W. Schul, W. Vermeulen, J.H. Hoeijmakers, R. van Driel, A.A. van Zeeland, L.H. Mullenders, Sequential assembly of the nucleotide excision repair factors in vivo, Mol. Cell 8 (2001) 213-224

[15] T. Riedl, F. Hanaoka, J.M. Egly, The comings and goings of nucleotide excision repair factors on damaged DNA, EMBO J. 22 (2003) 5293-5303

[16] R. Nishi, Y. Okuda, E. Watanabe, T. Mori, S. Iwai, C. Masutani, K. Sugasawa, F. Hanaoka, Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein, Mol. Cell. Biol. 25 (2005) 5664-5674

[17] D. Mu, M. Wakasugi, D.S. Hsu, A. Sancar, Characterization of reaction intermediates of human excision repair nuclease, J. Biol. Chem. 272 (1997) 28971-28979

[18] J.T. Reardon, A. Sancar, Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by the human excision nuclease, Genes Dev. 17 (2003) 2539-2551

[19] E. Evans, J.G. Moggs, J.R. Hwang, J.M. Egly, R.D. Wood, Mechanism of open complex and dual incision formation by human nucleotide excision repair factors, EMBO J. 16 (1997) 6559-6573

[20] E. Evans, J. Fellows, A. Coffer, R.D. Wood, Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein, EMBO J. 16 (1997) 625-638

[21] J.C. Huang, D.L. Svoboda, J.T. Reardon, A. Sancar, Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer, Proc. Natl. Acad. Sci. USA 89 (1992) 3664-3668

[22] A. O'Donovan, A.A. Davies, J.G. Moggs, S.C. West, R.D. Wood, XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair, Nature 371 (1994) 432-435

[23] A.M. Sijbers, W.L. de Laat, R.R. Ariza, M. Biggerstaff, Y.F. Wei, J.G. Moggs, K.C. Carter, B.K. Shell, E. Evans, M.C. de Jong, S. Rademakers, J. de Rooij, N.G. Jaspers, J.H. Hoeijmakers, R.D. Wood, Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease, Cell 86 (1996) 811-822

[24] T. Bessho, A. Sancar, L.H. Thompson, M.P. Thelen, Reconstitution of human excision nuclease with recombinant XPF-ERCC1 complex, J. Biol. Chem. 272 (1997) 3833-3837

[25] J.T. Reardon, P. Spielmann, J.C. Huang, S. Sastry, A. Sancar, J.E. Hearst, Removal of psoralen monoadducts and crosslinks by human cell free extracts, Nucleic Acids Res. 19 (1991) 4623-4629

[26] A.B. Houtsmuller, S. Rademakers, A.L. Nigg, D. Hoogstraten, J.H. Hoeijmakers, W. Vermeulen, Action of DNA repair endonuclease ERCC1/XPF in living cells, Science 284 (1999) 958-961

[27] D. Hoogstraten, A.L. Nigg, H. Heath, L.H. Mullenders, R. van Driel, J.H. Hoeijmakers, W. Vermeulen, A.B. Houtsmuller, Rapid switching of TFIIH between RNA polymerase I and II transcription and DNA repair in vivo, Mol. Cell 10 (2002) 1163-1174

[28] A. Politi, M.J. Mone, A.B. Houtsmuller, D. Hoogstraten, W. Vermeulen, R. Heinrich, R. van Driel, Mathematical modeling of nucleotide excision repair reveals efficiency of sequential assembly strategies, Mol. Cell 19 (2005) 679-690

[29] H. Naegeli, Mechanisms of DNA damage recognition in mammalian nucleotide excision repair, FASEB J. 9 (1995) 1043-1050

[30] K. Sugasawa, J.M. Ng, C. Masutani, S. Iwai, P.J. van der Spek, A.P. Eker, F. Hanaoka, D. Bootsma, J.H. Hoeijmakers, Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair, Mol. Cell 2 (1998) 223-232

[31] D. Batty, V. Rapic'-Otrin, A.S. Levine, R.D. Wood, Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged sites, J. Mol. Biol. 300 (2000) 275-290

[32] M. Wakasugi, A. Sancar, Order of assembly of human DNA repair excision nuclease, J. Biol. Chem. 274 (1999) 18759-18768

[33] J.T. Reardon, A. Sancar, Thermodynamic cooperativity and kinetic proofreading in DNA damage recognition and repair, Cell Cycle 3 (2004) 141-144

[34] K. Sugasawa, T. Okamoto, Y. Shimizu, C. Masutani, S. Iwai, F. Hanaoka, A multistep damage recognition mechanism for global genomic nucleotide excision repair, Genes Dev. 15 (2001) 507-521

[35] J.T. Reardon, A. Sancar, Molecular anatomy of the human excision nuclease assembled at sites of DNA damage, Mol. Cell. Biol. 22 (2002) 5938-5945

[36] R.S. Feldberg, L. Grossman, A DNA binding protein from human placenta specific for ultraviolet damaged DNA, Biochemistry 15 (1976) 2402-2408

[37] J. Moser, M. Volker, H. Kool, S. Alekseev, H. Vrieling, A. Yasui, A.A. van Zeeland, L.H. Mullenders, The UV-damaged DNA binding protein mediates efficient targeting of the nucleotide excision repair complex to UV-induced photo lesions, DNA Rep. 4 (2005) 571-582

[38] G. Kulaksiz, J.T. Reardon, A. Sancar, Xeroderma pigmentosum complementation group E protein (XPE/DDB2): purification of various complexes of XPE and analyses

of their damaged DNA binding and putative DNA repair properties, Mol. Cell. Biol. 25 (2005) 9784-9792

[39] M. Wakasugi, A. Kawashima, H. Morioka, S. Linn, A. Sancar, T. Mori, O. Nikaido, T. Matsunaga, DDB accumulates at DNA damage sites immediately after UV irradiation and directly stimulates nucleotide excision repair, J. Biol. Chem. 277 (2002) 1637-1640

[40] R. Dualan, T. Brody, S. Keeney, A.F. Nichols, A. Admon, S. Linn, Chromosomal localization and cDNA cloning of the genes (*DDB1* and *DDB2*) for the p127 and p48 subunits of a human damage-specific DNA binding protein, Genomics 29 (1995) 62-69

[41] A.F. Nichols, P. Ong, S. Linn, Mutations specific to the xeroderma pigmentosum group E $DDB^{(-)}$ phenotype, J. Biol. Chem. 271 (1996) 24317-24320

[42] V. Rapic-Otrin, V. Navazza, T. Nardo, E. Botta, M. McLenigan, D.C. Bisi, A.S. Levine, M. Stefanini, True XP group E patients have a defective UV-damaged DNA binding protein complex and mutations in *DDB2* which reveal the functional domains of its p48 product, Hum. Mol. Genet. 12 (2003) 1507-1522

[43] B.J. Hwang, G. Chu, Purification and characterization of a human protein that binds to damaged DNA, Biochemistry 32 (1993) 1657-1666

[44] J.T. Reardon, A.F. Nichols, S. Keeney, C.A. Smith, J.S. Taylor, S. Linn, A. Sancar, Comparative analysis of binding of human damaged DNA-binding protein (XPE) and *Escherichia coli* damage recognition protein (UvrA) to the major ultraviolet photoproducts: T[c,s]T, T[t,s]T, T[6-4]T, and T[Dewar]T, J. Biol. Chem. 268 (1993) 21301-21308

[45] B.O. Wittschieben, S. Iwai, R.D. Wood, DDB1-DDB2 (xeroderma pigmentosum group E) protein complex recognizes a cyclobutane pyrimidine dimer, mismatches,

apurinic/apyrimidinic sites, and compound lesions in DNA, J. Biol. Chem. 280 (2005) 39982-39989

[46] Y. Fujiwara, C. Masutani, T. Mizukoshi, J. Kondo, F. Hanaoka, S. Iwai, Characterization of DNA recognition by the human UV-damaged DNA-binding protein, J. Biol. Chem. 274 (1999) 20027-20033

[47] B.J. Hwang, S. Toering, U. Francke, G. Chu, p48 Activates a UV-damaged-DNA binding factor and is defective in xeroderma pigmentosum group E cells that lack binding activity, Mol. Cell. Biol. 18 (1998) 4391-4399

[48] J.Y. Tang, B.J. Hwang, J.M. Ford, P.C. Hanawalt, G. Chu, Xeroderma pigmentosum p48 gene enhances global genomic repair and suppresses UV-induced mutagenesis, Mol. Cell 5 (2000) 737-744

[49] M.E. Fitch, S. Nakajima, A. Yasui, J.M. Ford, In vivo recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product, J. Biol. Chem. 278 (2003) 46906-46910

[50] G. Chu, E. Chang, Xeroderma pigmentosum group E cells lack a nuclear factor that binds to damaged DNA, Science 242 (1988) 564-567

[51] A. Payne, G. Chu, Xeroderma pigmentosum group E binding factor recognizes a broad spectrum of DNA damage, Mutat. Res. 310 (1994) 89-102

[52] J.K. Kim, B.S. Choi, The solution structure of DNA duplex-decamer containing the (6-4) photoproduct of thymidylyl(3'→5')thymidine by NMR and relaxation matrix refinement, Eur. J. Biochem. 228 (1995) 849-854

[53] A.F. Nichols, T. Itoh, J.A. Graham, W. Liu, M. Yamaizumi, S. Linn, Human DNA damage-specific protein p48. Characterization of XPE mutations and regulation following UV irradiation, J. Biol. Chem. 275 (2000) 21422-21428

[54] J. Li, Q.E. Wang, Q. Zhu, M.A. El-Mahdy, G. Wani, M. Praetorius-Ibba, A.A. Wani, DNA damage binding protein component DDB1 participates in nucleotide excision repair through DDB2 DNA-binding and cullin 4A ubiquitin ligase activity, Cancer Res. 66 (2006) 8590-8597

[55] W. Liu, A.F. Nichols, J.A. Graham, R. Dualan, A. Abbas, S. Linn, Nuclear transport of human DDB protein induced by ultraviolet light. J. Biol. Chem, 275 (2000) 21429-21434

[56] P. Shiyanov, S.A. Hayes, M. Donepudi, A.F. Nichols, S. Linn, B.L Slagle, P. Raychaudhuri, The naturally occurring mutants of DDB are impaired in stimulating nuclear import of the p125 subunit and E2F1-activated transcription, Mol. Cell. Biol. 19 (1999) 4935-4943

[57] S. Angers, T. Li, X. Yi, M.J. MacCross, R.T. Moon, N. Zheng, Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery, Nature 443 (2006) 590-593

[58] P. Shiyanov, A. Nag, P. Raychaudhuri, Cullin 4A associates with the UV-damaged DNA-binding protein DDB, J. Biol. Chem. 274 (1999) 35309-35312

[59] R. Groisman, J. Polanowska, I. Kuraoka, J. Sawada, M. Saijo, R. Drapkin, A.F. Kisselev, K. Tanaka, Y. Nakatani, The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage, Cell 113 (2003) 357-367

[60] X. Chen, Y. Zhang, L. Douglas, P. Zhou, UV-damaged DNA-binding proteins are targets of CUL-4A-mediated ubiquitination and degradation, J. Biol. Chem. 276 (2001) 48175-48182

[61] A. Nag, T. Bondar, S. Shiv, P. Raychaudhuri, The xeroderma pigmentosum group E gene product DDB2 is a specific target of cullin 4A in mammalian cells, Mol. Cell. Biol. 21 (2001) 6738-6747

[62] K. Sugasawa, Y. Okuda, M. Saijo, R. Nishi, N. Matsuda, G. Chu, T. Mori, S. Iwai, K. Tanaka, K. Tanaka, F. Hanaoka, UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex, Cell 121 (2005) 387-400

[63] H. Wang, L. Zhai, J. Xu, H.Y. Joo, S. Jackson, H. Erdjument-Bromage, P. Tempst, Y. Xiong, Y. Zhang, Histone H3 and H4 ubiquitylation by the CUL4-DDB-ROC1 ubiquitin ligase facilitates cellular response to DNA damage, Mol. Cell 22 (2006) 383-394

[64] M.G. Kapetanaki, J. Guerrero-Santoro, D.C. Bisi, C.L. Hsieh, V. Rapic-Otrin, A.S. Levine, The DDB1-CUL4ADDB2 ubiquitin ligase is deficient in xeroderma pigmentosum group E and targets histone H2A at UV-damaged DNA sites, Proc. Natl. Acad. Sci. USA 103 (2006) 2588-2593

[65] V. Rapic-Otrin, M.P. McLenigan, D.C. Bisi, M. Gonzalez, A.S. Levine, Sequential binding of UV DNA damage binding factor and degradation of the p48 subunit as early events after UV irradiation, Nucleic Acid Res. 30 (2002) 2588-2598

[66] T. Itoh, T. Mori, H. Ohkubo, M. Yamaizumi, A newly identified patient with clinical xeroderma pigmentosum phenotype has a non-sense mutation in the DDB2 gene and incomplete repair in (6-4) photoproducts, J. Invest. Dermatol. 113 (1999) 251-257

[67] T. Itoh, D. Cado, R. Kamide, S. Linn, DDB2 gene disruption leads to skin tumors and resistance to apoptosis after exposure to ultraviolet light but not a chemical carcinogen, Proc. Natl. Acad. Sci. USA 101 (2004) 2052-2057

[68] T. Yoon, A. Chakrabortty, R. Franks, T. Valli, H. Kiyokawa, P. Raychaudhuri, Tumor-prone phenotype of the DDB2-deficient mice, Oncogene 24 (2005) 469-478

[69] P.C. Hanawalt, Transcription-coupled repair and human disease, Science 266 (1994) 1957-1958

[70] V.A. Bohr, C.A. Smith, D.S. Okumoto, P.C. Hanawalt, DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO cells is much more efficient than in the genome overall, Cell 40 (1985) 359-369

[71] H.D. Madhani, V.A. Bohr, P.C. Hanawalt, Differential DNA repair in transcriptionally active and inactive proto-oncogenes: c-able and c-mos, Cell 45 (1986) 417-423

[72] I. Mellon, G. Spivak, P.C. Hanawalt, Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene, Cell 51 (1987) 241-249

[73] J.R. Mitchell, J.H. Hoeijmakers, L.J. Niedernhofer, Divide and conquer: nucleotide excision repair battles cancer and ageing, Curr. Opin. Cell. Biol. 15 (2003) 232-240

[74] V.S. Kalogeraki, S. Tornaletti, P.C. Hanawalt, Transcription arrest at a lesion in the transcribed DNA strand in vitro is not affected by a nearby lesion in the opposite strand, J. Biol. Chem. 278 (2003) 19558-19564

[75] J.P. Laine, J.M. Egly, Initiation of DNA repair mediated by a stalled RNA polymerase IIO, EMBO J. 25 (2006) 387-397

[76] J. Venema, A. van Hoffen, V. Karcagi, A.T. Natarajan, A.A. van Zeeland, L.H. Mullenders, Xeroderma pigmentosum complementation group C cells remove pyrimidine dimers selectively from the transcribed strand of active genes, Mol. Cell. Biol. 11 (1991) 4128-4134

[77] A. van Hoffen, A.T. Natarajan, L.V. Mayne, A.A. van Zeeland, L.H. Mullenders, J. Venema, Deficient repair of the transcribed strand of active genes in Cockayne's syndrome cells, Nucl. Acids Res. 21 (1993) 5890-5895

[78] M.H. Wade, E.H. Chu, Effects of DNA damaging agents on cultured fibroblasts derived from patients with Cockayne syndrome, Mutat. Res. 59 (1979) 49-60

[79] J.O. Andressoo, J.H. Hoeijmakers, Transcription-coupled repair and premature ageing, Mutat. Res. 577 (2005) 179-194

[80] C. Troelstra, A. van Gool, J. de Wit, W. Vermeulen, D. Bootsma, J.H. Hoeijmakers, ERCC6, a member of the subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes, Cell 71 (1992) 939-953

[81] E. Citterio, V. Van Den Boom, G. Schnitzler, R. Kanaar, E. Bonte, R.E. Kingston, J.H. Hoeijmakers, W. Vermeulen, ATP-dependent chromatin remodeling by the Cockayne syndrome B DNA repair-transcription-coupling factor, Mol. Cell. Biol. 20 (2000) 7643-7653

[82] D.B. Bregman, R. Halaban, A.J. van Gool, K.A. Henning, E.C. Friedberg, S.L. Warren, UV-induced ubiquitination of RNA polymerase II: a novel modification deficient in Cockayne syndrome cells, Proc. Natl. Acad. Sci. USA 93 (1996) 11586-11590

[83] K.B. Lee, D. Wang, S.J. Lippard, P.A. Sharp, Transcription-coupled and DNA damage-dependent ubiquitination of RNA polymerase II in vitro, Proc. Natl. Acad. Sci. USA 99 (2002) 4239-4244

[84] L.J. Yang, H. Jiang, K.M. Rangel, RNA polymerase II stalled on a DNA template during transcription elongation is ubiquitinated and the ubiquitination facilitates displacement of the elongation complex, Int. J. Oncol. 22 (2003) 683-689

[85] K.B. Lee, P.A. Sharp, Transcription-dependent polyubiquitination of RNA polymerase II requires lysine 63 of ubiquitin, Biochemistry 43 (2004) 15223-15229

[86] A.H. Sarker, S.E. Tsutakawa, S. Kostek, C. Ng, D.S. Shin, M. Peris, E. Campeau, J.A. Tainer, E. Nogales, P.K. Cooper, Recognition of RNA polymerase II and

transcription bubbles by XPG, CSB, and TFIIH: insights for transcription-coupled repair and Cockayne Syndrome, Mol. Cell 20 (2005) 187-198

[87] P.C. Hanawalt, R.H. Haynes, Repair replication of DNA in bacteria: irrelevance of chemical nature of base defect, Biochem. Biophys. Res. Commun. 19 (1965) 462-467

[88] D. Gunz, M.T. Hess, H. Naegeli, Recognition of DNA adducts by human nucleotide excision repair. Evidence for a thermodynamic probing mechanism, J. Biol. Chem. 271 (1996) 25089-25098

[89] R. Muheim-Lenz, T. Buterin, G. Marra, H. Naegeli, Short-patch correction of C/C mismatches in human cells, Nucleic Acids Res. 32 (2004) 6696-6705

[90] T. Buterin, C. Meyer, B. Giese, H. Naegeli, DNA quality control by conformational readout on the undamaged strand of the double helix, Chem. Biol. 12 (2005) 913-922

[91] M.T. Hess, U. Schwitter, M. Petretta, B. Giese, H. Naegeli, Bipartite substrate discrimination by human nucleotide excision repair, Proc. Natl. Acad. Sci. USA 94 (1997) 6664-6669

[92] N. Buschta-Hedayat, T. Buterin, M.T. Hess, M. Missura, H. Naegeli, Recognition of nonhybridizing base pairs during nucleotide excision repair of DNA, Proc. Natl. Acad. Sci. USA 96 (1999) 6090-6095

[93] T. Buterin, M.T. Hess, N. Luneva, N.E. Geacintov, S. Amin, H. Kroth, A. Seidel, H. Naegeli, Unrepaired fjord region polycyclic aromatic hydrocarbon-DNA adducts in ras codon 61 mutational hot spots, Cancer Res. 60 (2000) 1849-1856

[94] R. Dip, U. Camenisch, H. Naegeli, Mechanisms of DNA damage recognition and strand discrimination in human nucleotide excision repair, DNA Repair 3 (2004) 1409-1423

[95] T.M. Hess, D. Gunz, N. Luneva, N.E. Geacintov, H. Naegeli, Base pair conformation-dependent excision of benzo[a]pyrene diol epoxide-guanine adducts by human nucleotide excision repair enzymes, Mol. Cell. Biol. 17 (1997) 7069-7076

[96] V. Mocquet, K. Kropachev, M. Kolbanovskiy, A. Kolbanovskiy, A. Tapias, Y. Cai, S. Broyde, N.E. Geacintov, J.M. Egly, The human DNA repair factor XPC-HR23B distinguishes stereoisomeric benzo[a]pyrenyl-DNA lesions, EMBO J. 26 (2007) 2923-2932

[97] J.C. Huang, D.S. Hsu, A. Kazantsev, A. Sancar, Substrate spectrum of human excinuclease: repair of abasic sites, methylated bases, mismatches and bulky adducts, Proc. Natl. Acad. Sci. USA 91 (1994) 12213-12217

[98] B. Van Houten, D.L. Croteau, M.J. DellaVecchia, H. Wang, C. Kisker, Close-fitting sleeves: DNA damage recognition by the UvrABC nuclease system, Mutat. Res. 577 (2005) 92-117

[99] J.B. Mills, P.J. Hagerman, Origin of the intrinsic rigidity of DNA, Nucleic Acids Res. 32 (2004) 4055-4059

[100] R.J. Isaacs, H.P. Spielmann, A model for initial DNA lesion recognition by NER and MMR based on local conformational flexibility, DNA Repair 3 (2004) 455-464

[101] W. Yang, Poor base stacking at DNA lesions may initiate recognition by many repair proteins, DNA Repair 5 (2006) 654-666

[102] C.H. Choi, G. Kalosakas, K.O. Rasmussen, M. Hiromura, A.R. Bishop, A. Usheva, DNA dynamically directs its own transcription initiation, Nucl. Acids. Res. 32 (2004) 1584-1590

[103] T.S. van Erp, S. Cuesta-Lopez, J.G. Hagmann, M. Peyrard, Can one predict DNA transcription start sites by studying bubbles?, Phys. Rev. Lett. 95 (2005) 218104

[104] B.S. Alexandrov, L.T. Wille, K.O. Rasmussen, A.R. Bishop, K.B. Blagoev, Bubble statistics and dynamics in double-stranded DNA, Phys. Rev. E. 74 (2006) 050901

[105] K.B. Blagoev, B.S. Alexandrov, E.H. Goodwin, A.R. Bishop, Ultra-violet light induced changes in DNA dynamics may enhance TT-dimer recognition, DNA Repair 5 (2006) 863-867

[106] C. Masutani, K. Sugasawa, J. Yanagisawa, T. Sonoyama, M. Ui, T. Enomoto, K. Takio, K. Tanaka, P.J. van der Spek, D. Bootsma, J.H. Hoeijmakers, F. Hanaoka, Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23, EMBO J. 13 (1994) 1831-1843

[107] J.M. Ng, W. Vermeulen, G.T. van der Horst, S. Bergink, K. Sugasawa, H. Vrieling, J.H. Hoeijmakers, A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein, Genes Dev. 17 (2003) 1630-1645

[108] R. Kusumoto, C. Masutani, K. Sugasawa, S. Iwai, M. Araki, A. Uchida, T. Mizukoshi, F. Hanaoka, Diversity of the damage recognition step in the global genomic nucleotide excision repair in vitro, Mutat. Res. 485 (2001) 219-227

[109] K.S. Trego, J.J. Turchi, Pre-steady-state binding of damaged DNA by XPC-hHR23B reveals a kinetic mechanism for damage discrimination, Biochemistry 45 (2006) 1961-1969

[110] A. Janicijevic, K. Sugasawa, Y. Shimizu, F. Hanaoka, N. Wijgers, M. Djurica, J.H. Hoeijmakers, C. Wyman, DNA bending by the human damage recognition complex XPC-HR23B, DNA Repair 2 (2003) 325-336

[111] K. Sugasawa, Y. Shimizu, S. Iwai, F. Hanaoka, A molecular mechanism for DNA damage recognition by the xeroderma pigmentosum group C protein complex, DNA Repair 1 (2002) 95-107

[112] L. Li, X. Lu, C. Peterson, R. Legerski, XPC interacts with both HHR23B and HHR23A in vivo, Mutat. Res. 383 (1997) 197-203

[113] A. Uchida, K. Sugasawa, C. Masutani, N. Dohmae, M. Araki, M. Yokoi, Y. Ohkuma, F. Hanaoka, The carboxy-terminal domain of the XPC protein plays a crucial role in nucleotide excision repair through interactions with transcription factor IIH, DNA Repair 1 (2002) 449-461

[114] C.G. Bunick, M.R. Miller, B.E. Fuller, E. Fanning, W.J. Chazin, Biochemical and structural domain analysis of xeroderma pigmentosum complementation group C protein, Biochemistry 45 (2006) 14965-14979

[115] J.S. You, M. Wang, S.H. Lee, Biochemical analysis of the damage recognition process in nucleotide excision repair, J. Biol. Chem. 278 (2003) 7476-7485

[116] M. Wakasugi, A. Sancar, Assembly, subunit composition, and footprint of human DNA repair excision nuclease, Proc. Natl. Acad. Sci. USA 95 (1998) 6669-6674

[117] V. Anantharaman, E.V. Koonin, L. Aravind, Peptide-N-glycanases and DNA repair proteins, Xp-C/Rad4, are, respectively, active and inactivated enzymes sharing a common transglutaminase fold, Hum. Mol. Genet. 10 (2001) 1627-1630

[118] O. Maillard, S. Solyom, H. Naegeli, An aromatic sensor with aversion to damaged strands confers versatility to DNA repair, PLoS Biol. 5 (2007) e79

[119] A. Bochkarev, E. Bochkareva, From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold, Curr. Opin. Struct. Biol. 14 (2004) 36-42

[120] F. Miao, M. Bouziane, R. Dammann, C. Masutani, F. Hanaoka, G. Pfeifer, T.R. O'Connor, 3-Methyladenine-DNA glycosylase (MPG protein) interacts with human RAD23 proteins, J. Biol. Chem. 275 (2000) 28433-28438

[121] Y. Shimizu, S. Iwai, F. Hanaoka, K. Sugasawa, Xeroderma pigmentosum group C protein interacts physically and functionally with thymine DNA glycosylase, EMBO J. 22 (2003) 164-173

[122] L.M. Iakoucheva, A.L. Kimzey, C.D. Masselon, J.E. Bruce, E.C. Garner, C.J. Brown, A.K. Dunker, R.D. Smith, E.J. Ackerman, Identification of intrinsic order and disorder in the DNA repair protein XPA, Protein Sci. 10 (2001) 560-571

[123] L.M. Iakoucheva, A.L. Kimzey, C.D. Masselon, R.D. Smith, A.K. Dunker, E.J. Ackerman, Aberrant mobility phenomena of the DNA repair protein XPA, Protein Sci. 10 (2001) 1353-1362

[124] T. Ikegami, I. Kuraoka, M. Saijo, N. Kodo, Y. Kyogoku, K. Morikawa, K. Tanaka, M. Shirakawa, Resonance assignments, solution structure, and backbone dynamics of the DNA- and RPA-binding domain of human repair factor XPA, J. Biochem. 125 (1999) 495-506

[125] Z.G. Yang, Y. Liu, L.Y. Mao, J.T. Zhang, Y. Zou, Dimerization of human XPA and formation of XPA2-RPA protein complex, Biochemistry 41 (2002) 13012-13020

[126] S. Rademakers, M. Volker, D. Hoogstraten, A.L. Nigg, M.J. Mone, A.A. Van Zeeland, J.H. Hoeijmakers, A.B. Houtsmuller, W. Vermeulen, Xeroderma pigmentosum group A protein loads as a separate factor onto DNA lesions, Mol. Cell. Biol. 23 (2003) 5755-5767

[127] L. Li, X. Lu, C.A. Peterson, R.J. Legerski, An interaction between the DNA repair factor XPA and replication protein A appears essential for nucleotide excision repair, Mol. Cell. Biol. 15 (1995) 5396-5402

[128] C.H. Park, D. Mu, J.T. Reardon, A. Sancar, The general transcription-repair factor TFIIH is recruited to the excision repair complex by the XPA protein independent of the TFIIE transcription factor, J. Biol. Chem. 270 (1995) 4896-4902

[129] L. Li, S.J. Elledge, C.A. Peterson, E.S. Bales, R.J. Legerski, Specific association between the human DNA repair proteins XPA and ERCC1, Proc. Natl. Acad. Sci. USA 91 (1994) 5012-5016

[130] C.H. Park, A. Sancar, Formation of a ternary complex by human XPA, ERCC1, and ERCC4(XPF) excision repair proteins, Proc. Natl. Acad. Sci. USA 91 (1994) 5017-5021

[131] L. Li, C.A. Peterson, X. Lu, R.J. Legerski, Mutations in XPA that prevent association with ERCC1 are defective in nucleotide excision repair, Mol. Cell. Biol. 15 (1995) 1993-1998

[132] P. Robins, C.J. Jones, M. Biggerstaff, T. Lindahl, R.D. Wood, Complementation of DNA repair in xeroderma pigmentosum group A cell extracts by a protein with affinity for damaged DNA, EMBO J. 10 (1991) 3913-3921

[133] C.J. Jones, R.D. Wood, Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA, Biochemistry 32 (1993) 12096-12104

[134] Z. He, L.A. Henricksen, M.S. Wold, C.J. Ingles, RPA involvement in the damage-recognition and incision steps of nucleotide excision repair, Nature 374 (1995) 566-569

[135] M. Missura, T. Buterin, R. Hindges, U. Hubscher, J. Kasparkova, V. Brabec, H. Naegeli, Double-check probing of DNA bending and unwinding by XPA-RPA: an architectural function in DNA repair, EMBO J. 20 (2001) 3554-3564

[136] S.M. Patrick, J.J. Turchi, Xeroderma pigmentosum complementation group A protein (XPA) modulates RPA-DNA interactions via enhanced complex stability and inhibition of strand separation activity, J. Biol. Chem. 277 (2002) 16096-16101

[137] A. Nagai, M. Saijo, I. Kuraoka, T. Matsuda, N. Kodo, Y. Nakatsu, T. Mimaki, M. Mino, M. Biggerstaff, R.D. Wood, A. Sijbers, J.H. Hoeijmakers, K. Tanaka, Enhancement of damage-specific DNA binding of XPA by interaction with the ERCC1 DNA repair protein, Biochem. Biophys. Res. Commun. 211 (1995) 960-966

[138] U. Camenisch, R. Dip, S.B. Schumacher, B. Schuler, H. Naegeli, Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair, Nat. Struct. Mol. Biol. 13 (2006) 278-284

[139] M.H. Werner, A.M. Gronenborn, G.M. Clore, Intercalation, DNA kinking, and the control of transcription, Science 271 (1996) 778-784

[140] Q. Shi, R. Thresher, A. Sancar, J. Griffith, Electron microscopic study of (A)BC excinuclease, DNA is sharply bent in the UvrB-DNA complex, J. Mol. Biol. 226 (1992) 425-432

[141] T. Ikegami, I. Kuraoka, M. Saijo, N. Kodo, Y. Kyogoku, K. Morikawa, K. Tanaka, M. Shirakawa, Solution structure of the DNA- and RPA-binding domain of the human repair factor XPA, Nat. Struct. Biol. 5 (1998) 701-706

[142] G.W. Buchko, S. Ni, B.D. Thrall, M.A. Kennedy, Structural features of the minimal DNA binding domain (M98-F219) of human nucleotide excision repair protein XPA, Nucleic Acids Res. 26 (1998) 2779-2788

[143] G.W. Buchko, G.W. Daughdrill, R. de Lorimier, B.K. Rao, N.G. Isern, J.M. Lingbeck, J.S. Taylor, M.S. Wold, M. Gochin, L.D. Spicer, D.F. Lowry, M.A. Kennedy, Interactions of human nucleotide excision repair protein XPA with DNA

and RPA70 Delta C327: chemical shift mapping and 15N NMR relaxation studies, Biochemistry 38 (1999) 15116-15128

[144] U. Camenisch, R. Dip, M. Vitanescu, H. Naegeli, Xeroderma pigmentosum complementation group A protein is driven to nucleotide excision repair sites by the electrostatic potential of distorted DNA, DNA Rep., in press.

[145] D.G. Vassylyev, T. Kashiwagi, Y. Mikami, M. Ariyoshi, S. Iwai, E. Ohtsuka, K. Morikawa, Atomic model of a pyrimidine dimer excision repair enzyme complexed with a DNA substrate: structural basis for damaged DNA recognition, Cell 83 (1995) 773-782

[146] P. Schultz, S. Fribourg, A. Poterszman, V. Mallouh, D. Moras, J.M. Egly, Molecular structure of human TFIIH, Cell 102 (2000) 599-607

[147] W.H. Chang, R.D. Kornberg, Electron crystal structure of the transcription factor and DNA repair complex, core TFIIH, Cell 102 (2000) 609-613

[148] F. Coin, L. Proietti De Santis, T. Nardo, O. Zlobinskaya, M. Stefanini, J.M. Egly, p8/TTD-A as a repair-specific TFIIH subunit, Mol. Cell 21 (2006) 215-226

[149] F.C. Holstege, P.C. van der Vliet, H.T. Timmers, Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH, EMBO J. 15 (1996) 1666-1677

[150] M. Yokoi, C. Masutani, T. Maekawa, K. Sugasawa, Y. Ohkuma, F. Hanaoka, The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIH to damaged DNA, J. Biol. Chem. 275 (2000) 9870-9875

[151] S.N. Guzder, H. Qiu, C.H. Sommers, P. Sung, L. Prakash, S. Prakash, DNA repair gene RAD3 of S. cerevisiae is essential for transcription by RNA polymerase II, Nature 367 (1994) 91-94

[152] M. Zurita, C. Merino, The transcriptional complexity of the TFIIH complex, Trends Genet. 19 (2003) 578-584

[153] M.S. Dillingham, M. Spies, S.C. Kowalczykowski, RecBCD enzyme is a bipolar DNA helicase, Nature 423 (2003) 893-897

[154] F. Coin, J.C. Marinoni, C. Rodolfo, S. Fribourg, A.M. Pedrini, J.M. Egly, Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH, Nat. Genet. 20 (1998) 184-188

[155] R.J. Bienstock, M. Skorvaga, B.S. Mandavilli, B. Van Houten, Structural and functional characterization of the human DNA repair helicase XPD by comparative molecular modeling and site-directed mutagenesis of the bacterial repair protein UvrB, J. Biol. Chem. 278 (2003) 5309-5316

[156] K. Theis, P. Chen, M. Skorvaga, B. Van Houten, C. Kisker, Crystal structure of UvrB, a DNA helicase adapted for nucleotide excision repair, EMBO J. 18 (1999) 6899-6907

[157] M. Machius, L. Henry, M. Palnitkar, J. Deisenhofer, Crystal structure of the DNA nucleotide excision repair enzyme UvrB from Thermus thermophilus, Proc. Natl. Acad. Sci. USA 96 (1999) 11717-11722

[158] N. Nakagawa, M. Sugahara, R. Masui, R. Kato, K. Fukuyama, S. Kuramitsu, Crystal structure of Thermus thermophilus HB8 UvrB protein, a key enzyme of nucleotide excision repair, J. Biochem. 126 (1999) 986-990

[159] J.J. Truglio, E. Karakas, B. Rhau, H. Wang, M.J. DellaVecchia, B. Van Houten, C. Kisker, Structural basis for DNA recognition and processing by UvrB, Nat. Struct. Mol. Biol. 13 (2006) 360-364

[160] H. Naegeli, L. Bardwell, E.C. Friedberg, The DNA helicase and adenosine triphosphatase activities of yeast Rad3 protein are inhibited by DNA damage, J. Biol. Chem. 267 (1992) 392-398

[161] H. Naegeli, L. Bardwell, E.C. Friedberg, Inhibition of Rad3 DNA helicase activity by DNA adducts and abasic sites: implications for the role of a DNA helicase in damage-specific incision of DNA, Biochemistry 32 (1993) 613-621.

[162] M.S. Wold, Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism, Annu. Rev. Biochem. 66 (1997) 61-92

[163] W.L. de Laat, E. Appeldoorn, K. Sugasawa, E. Weterings, N.G. Jaspers, J.H. Hoeijmakers, DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair, Genes Dev. 12 (1998) 2598-2609

[164] I.L. Hermanson-Miller, J.J. Turchi, Strand-specific binding of RPA and XPA to damaged duplex DNA, Biochemistry 41 (2002) 2402-2408

[165] J.H. Lee, C.J. Park, A.I. Arunkumar, W.J. Chazin, B.S. Choi, NMR study on the interaction between RPA and DNA decamer containing cis-syn cyclobutane pyrimidine dimer in the presence of XPA: implication for damage verification and strand-specific dual incision in nucleotide excision repair, Nucleic Acid Res. 31 (2003) 4747-4754

[166] Y. Liu, Y. Liu, Z. Yang, C. Utzat, G. Wang, A.K. Basu, Y. Zou, Cooperative interaction of human XPA stabilizes and enhances specific binding of XPA to DNA damage, Biochemistry 44 (2005) 7361-7368

[167] J.J. Lin, A. Sancar, (A)BC excinuclease: the Escherichia coli nucleotide excision repair enzyme, Mol. Microbiol. 6 (1992) 2219-2224

[168] M.S. Junop, G. Obmolova, K. Rausch, P. Hsieh, W. Yang, Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair, Mol. Cell 7 (2001) 1-12

Chapter IV

DNA Repair Triggered by Sensors of Helical Dynamics

DNA repair triggered by sensors of helical dynamics

Olivier Maillard[1], Ulrike Camenisch[1], Flurina C. Clement[1], Krastan B. Blagoev[2] and Hanspeter Naegeli[1]

[1]Institute of Pharmacology and Toxicology, University of Zürich-Vetsuisse, Winterthurerstrasse 260, CH-8057 Zürich, Switzerland

[2]Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

Corresponding author: Naegeli, H. (naegelih@vetpharm.unizh.ch)

Ref.: Ms. No. TIBS-D-07-00027R2
Trends in Biochemical Sciences
Accepted, 29. August 2007

Nucleotide excision repair is a constitutive stress response that eliminates DNA lesions induced by multiple genotoxic agents. Unlike the immune system, which generates billions of immunoglobulins and T-cell receptors for antigen recognition, the nucleotide excision repair complex uses only a few generic factors to detect an astounding diversity of DNA modifications. New data favor an unexpected strategy whereby damage recognition is initiated by the detection of abnormal oscillations in the undamaged strand opposite to DNA lesions. Another core subunit recognizes the increased susceptibility of DNA to be kinked at injured sites. Thus, we suggest that early nucleotide excision repair factors gain substrate versatility by avoiding direct contacts with modified residues and exploiting, instead, the altered dynamics of damaged DNA duplexes.

Introduction: versatile DNA damage recognition

The genome is under permanent attack from ultraviolet (UV) radiation, environmental mutagens and endogenous metabolic byproducts that pose a continuous risk to DNA integrity. Given that genome stability represents a fundamental biological challenge, it is not surprising to find intricate networks of repair mechanisms that are able to cope with these manifold genetic insults. In particular, living organisms of all biological kingdoms are equipped with a multi-protein system that eliminates crosslinks between adjacent bases, primarily cyclobutane pyrimidine dimers and (6-4) photoproducts, induced by exposure to the UV component of sunlight [1]. This repair pathway, generally known as nucleotide excision repair (NER) is, in fact, the only cellular mechanism available for the error-free removal of such DNA photoproducts in humans. In addition to intrastrand crosslinks, the same NER pathway also eliminates a wide diversity of base adducts caused by electrophilic chemicals and oxygen radicals [2,3], as well as protein-DNA crosslinks [4], methylated bases and abasic sites [5]. These different lesions share no common chemical motif that may support a classic "lock and key" recognition scheme, raising the question of how NER systems achieve their characteristic substrate versatility.

The "cut and patch" NER pathway can be broken down into the following steps: (i) recognition of DNA distortions by XPC protein, (ii) verification of DNA lesions by TFIIH, XPA and RPA, (iii) dual DNA incision and (iv) DNA repair synthesis (Box 1). It has been demonstrated that four subunits (XPC, TFIIH, XPA and RPA) are necessary and sufficient to form a preincision complex that detects target sites for mammalian NER activity [6,7]. The mechanism by which the few subunits of this versatile recognition intermediate discriminate between damaged substrates and normal double-stranded DNA is a very actively investigated area. There is also ongoing discussion regarding the role of an auxiliary factor (UV-DDB) in facilitating the excision of cyclobutane pyrimidine dimers, which represent the most

abundant type of UV lesion [7,8]. Recent developments in the characterization of XPC protein indicate that this early recognition subunit of the NER pathway operates in an indirect manner by acting as a molecular sensor of the increased dynamics of damaged DNA compared to the more rigid native double helix. In view of these new findings, we propose a counterintuitive hypothesis by which the rapid recruitment of NER factors to DNA damage in mammalian chromosomes is triggered by abnormal strand oscillations or other conformational fluctuations of the undamaged complementary sequence across lesion sites.

Preassembly, stepwise recruitment or random aggregation?

The order of arrival of NER factors has been a matter of intense debate [7,9-11], in part raised by the fact that none of the core proteins seems to display a high enough selectivity to function as a unique sensor of damaged substrates. A full incision complex comprising 20 polypeptides achieves a mass of >1 MDa, but the measurement of nuclear diffusion rates indicates that NER factors are present as separate subunits and not as part of a large repair "machine". Also, the fast translocation of these factors to hotspots of DNA damage is more consistent with a sequential recruitment of individual subunits [12]. This stepwise assembly has the advantage that multifunctional components such as TFIIH, RPA or XPF-ERCC1 can shuttle between transcription, replication, recombination or other parallel pathways. In addition, the sequential recruitment of freely diffusing proteins may be more rapid and efficient than other strategies such as random aggregation or the preassembly of large "repairosome" complexes [12].

The "XPC first" model is the most widely accepted scenario for the initiation of NER activity (Box 1). This model was instigated by reconstitution experiments aimed at determining the order in which NER proteins are recruited to lesion sites. In one study, damaged plasmids pre-incubated with XPC were repaired more rapidly in cell extracts than

those pre-incubated with XPA and RPA [13]. Further support for the "XPC first" model came from the energy dependence of specific NER intermediates [10,11]. ATP is not necessary for the simple association of XPC and TFIIH with damaged DNA fragments, but ATP hydrolysis is absolutely required for the recruitment of XPA and RPA. These observations indicate that the damage-specific binding of XPC is an early ATP-independent step and that DNA unwinding promoted by TFIIH, an energy-dependent process, is needed for the subsequent incorporation of the remaining subunits into the growing NER complex.

The order of assembly has also been examined by monitoring the nuclear trafficking of core NER subunits. For that purpose, cultured human cells were exposed to UV light through filters with small pores to induce localized foci of DNA damage. The mobility of XPC and XPA from unirradiated regions of the nuclei to these damaged foci was monitored by staining each factor with fluorescently tagged antibodies [9]. This approach established that XPC accumulates in DNA repair foci even in the absence of XPA. Instead, XPA, TFIIH, XPG (the 3' endonuclease) and XPF-ERCC1 do not appear to move to damaged foci when XPC is missing, implying that the lesions have to be recognized by XPC before other subunits can be loaded onto the DNA substrate. Taken together, these different findings converge on the idea that the XPC subunit is the primary initiator of NER assembly.

An attraction for the undamaged side of the double helix

What characteristic feature of damaged DNA induces the recruitment of XPC protein? Hanawalt and Haynes [14] were the first to propose that the need for DNA repair is determined by comparing the conformation of damaged DNA to that of a normal double helix. It was later found that the recognition of different bulky base adducts correlates with the degree by which the duplex is thermodynamically destabilized [15]. More recently, Isaacs and Spielmann [16] concluded that damaged DNA is more easily bent and, therefore,

postulated that a higher local flexibility may provide another generic determinant shared by all known NER substrates.

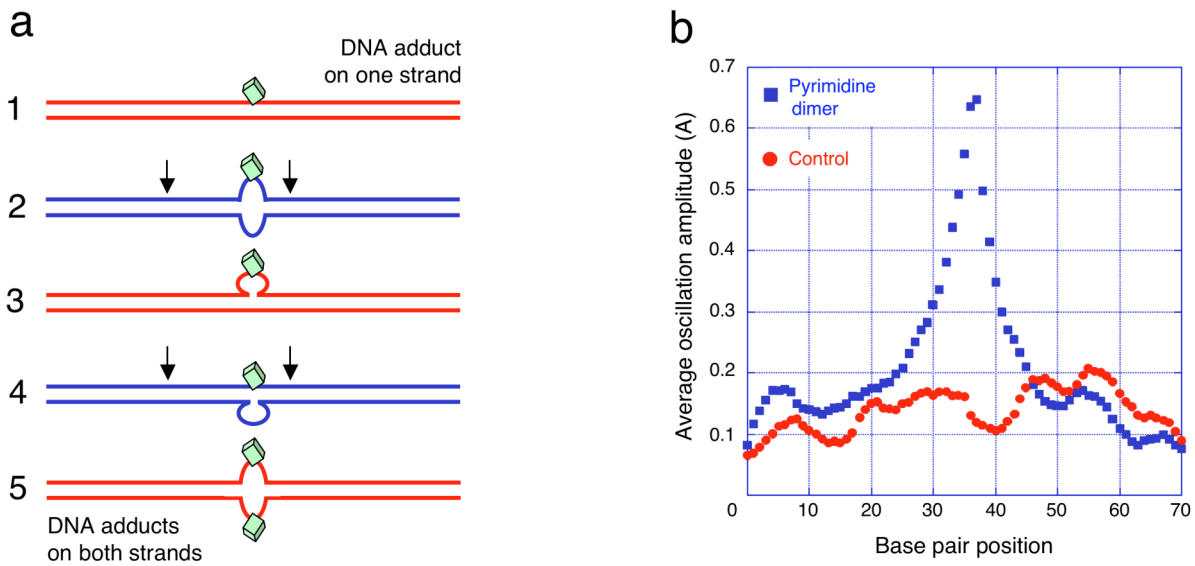


Figure 1 Recruitment of the NER complex to abnormally oscillating sites on the undamaged side of the DNA duplex. (**a**) Conformational requirement for excision of a "non-distorting" DNA adduct located in one strand of the duplex (substrates 1-4). In the case of substrate 5, the same DNA adduct was placed in both strands simultaneously. Substrates 1, 3 and 5 are refractory to recognition by the NER complex. Only the adducts in substrates 2 and 4 are conveyed to the NER process. The sites of dual DNA incision are indicated by the arrows. These experimental findings [18] indicate that DNA damage recognition is initiated by the appearance of a bulge of native but single-stranded DNA opposite to the offending lesion. (**b**) Strand oscillations caused by a single bipyrimidine photoproduct. The average amplitude of strand openings is shown at different positions of a 70-base pair fragment. Red circles, undamaged sequence; blue rectangles, DNA containing a cyclobutane pyrimidine dimer (adapted from Ref. 21).

The preference for lesions that lower the melting temperature of duplex DNA has been exploited to test the particular orientation with which the NER subunits align themselves along the double-stranded nucleic acid molecule. For that purpose, highly defined DNA substrates were constructed by combining an artificial adduct, known to retain normal hydrogen bonding and stacking interactions [17], with different kinds of conformational distortions. The varying repairability of these synthetic duplexes in human cell extracts confirmed that destabilization of the double helix is a crucial molecular determinant for lesion recognition. In fact, excision was readily detected on composite substrates in which the "non-distorting" adduct was combined with a DNA bulge generated by the insertion of mismatches or nucleotide loops (Figure 1a). This local disruption of base pairing led to excision regardless of whether the bulge included both strands of the duplex or only the undamaged complementary strand across the adduct (Figure 1a, substrates 2 and 4). More surprising, however, is the observation that no excision took place when the same adduct was combined with a one-sided bulge involving only the damaged strand (Figure 1a, substrate 3). Similarly, the duplex became refractory to excision when both strands at the site of distortion carried a bulky modification (Figure 1a, substrate 5). Thus, this novel approach not only confirmed that a local DNA bulge is indispensable for NER activity but, additionally, pointed to an unexpected mechanism in which the early sensing of DNA distortions occurs through an interaction with flipped-out nucleotides on the undamaged side of the double helix [18]. This goes against the conventional dogma asserting that DNA lesion recognition occurs primarily through contacts with damaged bases or at least some components of the damaged strand. Instead, the experiments of Figure 1a indicate that mammalian NER factors are initially loaded onto the opposite undamaged sequence of the double helix, from where the lesion is approached and subsequently repaired. Additional studies showed that the

recognition of extruded bases in the undamaged complementary strand is also crucial for the analogous NER process in bacteria [19].

Damage-induced DNA oscillations

The model of conformational "readout" on the undamaged side of the double helix, outlined in the previous section, raises the question of whether naturally arising DNA lesions may indeed destabilize the double helix to a sufficient degree to induce a local bulge with single-stranded character. New insights came from mathematical models that describe the complex macromolecular dynamics of modified nucleic acid substrates. In fact, double-stranded DNA is not a static entity and even the native double helix is constantly in motion due to thermal fluctuations. Normally, the two strands of duplex DNA oscillate with respect to each other in a sequence-dependent manner, such that the distance between complementary strands exhibits fast and small variations [20]. In the absence of DNA damage, however, the picosecond-to-nanosecond timescale of these spontaneous vibrations is too short to be recognized by repair factors. In contrast, normal base pairing and stacking interactions are weakened in the presence of DNA adducts, giving rise to much more prominent strand movements.

Recent mathematical calculations led to the prediction that even subtly distorting lesions, such as cyclobutane pyrimidine dimers, provoke longer-lived and larger openings between the two strands of the double helix relative to undamaged DNA [21]. The frequency of these large strand fluctuations is 25 times higher at photoproduct positions than in undamaged DNA sequences and, as illustrated in Figure 1b, the amplitude of these oscillations is drastically increased relative to the undamaged control. Another important prediction is that these dynamic changes triggered by base damage are expected to involve mainly oscillations

of the intact sequence opposing the lesion, as the strand containing base adducts is less flexible than native DNA [21].

To summarize, the paradigm of bipyrimidine photoproducts shows that DNA lesions have the ability to induce dynamic changes of the duplex, involving transient openings between complementary sequences, which result in strand oscillations particularly on the undamaged side of the double helix. The increased appearance of oscillations in the undamaged strand is consistent with a marked conformational heterogeneity observed across benzo[a]pyrene diol epoxide adducts [22]. In duplexes bearing an acetylaminofluorene adduct, only the undamaged sequence is susceptible to incision by endonuclease VII [23], an enzyme that cleaves distorted DNA, thus supporting the conclusion that more extensive dynamic changes occur in the complementary native strand.

XPC protein is a sensor of abnormal strand oscillations

One of the initiators of the mammalian NER reaction, XPC protein, has an intrinsic affinity for DNA molecules that deviate from the Watson-Crick geometry, including loops, mismatched bubbles or overhangs [24,25]. Accordingly, it has been suggested that XPC protein displays a general preference for conformational anomalies of the double helix, but the molecular basis of this unique recognition function has long remained elusive.

The human XPC sequence comprises a region of homology with the nucleic acid-binding domains of RPA and breast cancer susceptibility protein 2 (BRCA2) [26]. In the case of RPA and BRCA2, the single-stranded DNA-binding activity correlates with the presence of conserved aromatic side chains that mediate stacking interactions with closely spaced DNA bases [27]. The search for functionally analogous residues in human XPC protein revealed that two neighboring aromatics are essential for DNA-binding and NER activity [26]. XPC protein also displays a marked preference for single-stranded oligonucleotides, implying that

it recognizes the local single-stranded character of damaged DNA. However, the presence of DNA lesions strongly reduces its binding to damaged single-stranded oligonucleotides compared to the more efficient interaction of XPC with undamaged counterparts [26,28].

In summary, the affinity for single-stranded oligonucleotides, in combination with an aversion to interact with damaged strands, indicates that XPC protein initiates the NER pathway by "reading" the native strand of damaged duplexes. As illustrated in Figure 2, this mechanism fits with the large and long-lived oscillations across lesion sites, thus predicting that XPC protein operates by capturing the transient formation of single-stranded conformations in the complementary strand. At first glance, the suggestion that a DNA damage sensor is attracted by the undamaged complementary strand might seem internally inconsistent. However, this newly discovered mechanism of DNA quality control has the advantage that the initial sensor does not rely on interactions with damaged bases, and in fact avoids such intimate contacts with abnormal residues, thus enhancing the range of modifications that can be conveyed to DNA repair. In this respect, XPC may use its affinity for oscillating DNA strands to serve as a more general platform that attracts the components of multiple repair pathways to damaged sites in chromatin. For example, recent data suggest that XPC protein may interact with subunits of the base excision repair [29,30] and non-homologous end-joining systems [31].

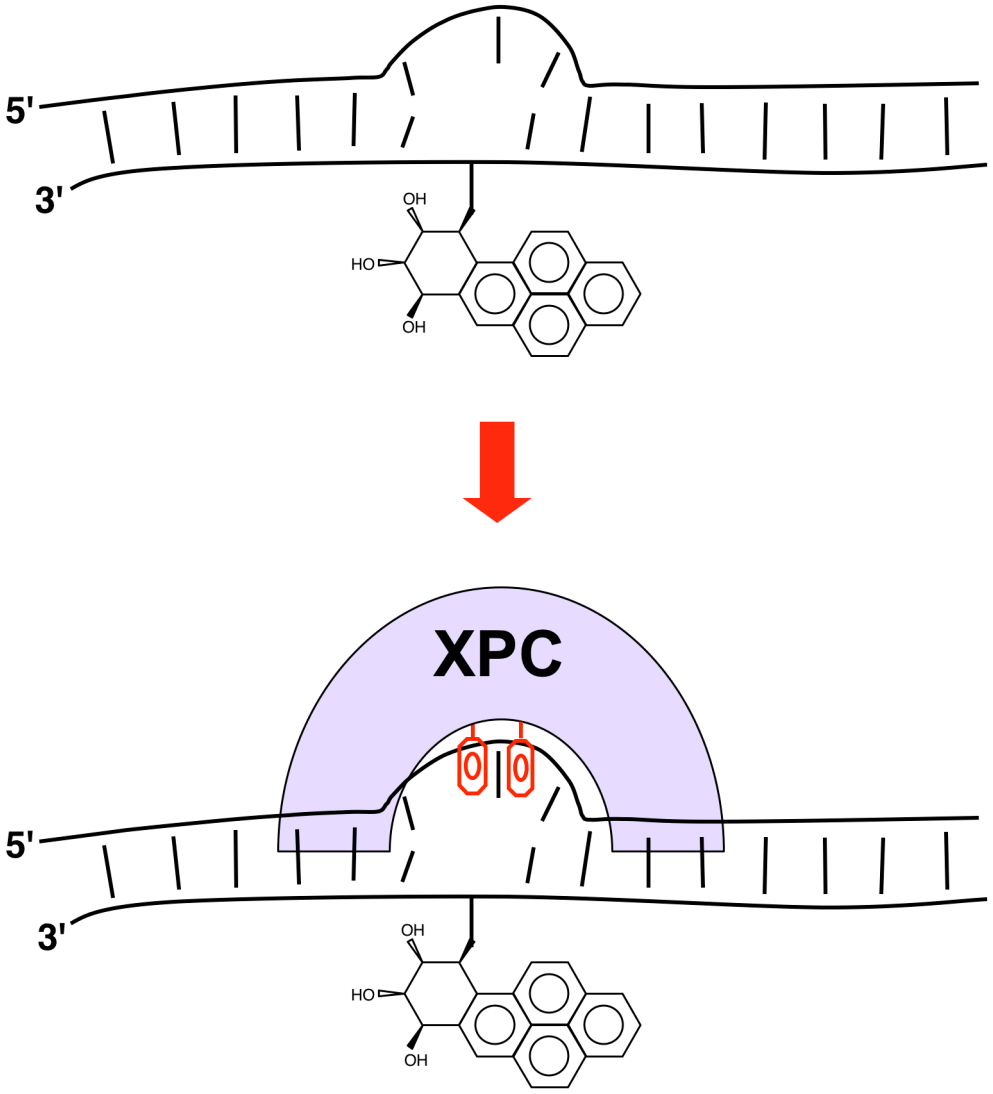


Figure 2 Scheme illustrating how human XPC protein may deploy a pair of aromatic side chains (shown in red color) to capture DNA oscillations in the undamaged strand opposite to base lesions (adapted from Ref. 26). This model predicts that unpaired deoxyribonucleotide residues located across the damaged site become sandwiched between aromatic side chains of XPC protein. The substrate versatility of this unique sensor of DNA oscillations is enhanced by avoiding direct contacts with adducted or otherwise defective residues in the damaged strand.

The mode of action of other NER recognition factors

Due to the deleterious consequences of strand breaks, the incision reaction must occur under an extremely rigorous control. Lin and Sancar [32] predicted that a cascade of different DNA damage recognition steps of mediocre selectivity leads to an overall specificity that is comparable to that of transcription factors. For example, XPC protein, which interacts mainly with the undamaged strand, mediates the loading of TFIIH onto the damaged strand [33], such that the two DNA helicase components of TFIIH (XPB and XPD) function as tracking enzymes that locate the chemically damaged base [34,35].

XPA is another subunit of the ultimate lesion verification complex that, like XPC, has a preference for distorted DNA structures carrying mismatches, loops or bubbles, even if no actual DNA lesion has been introduced into the substrate [36]. Despite its low DNA-binding affinity, a damage "verification" function has been proposed for XPA based on the observation that it interacts more avidly with DNA in conjunction with other NER partners [37]. XPA exhibits a distinctive preference for distorted DNA molecules, such as cisplatin-damaged duplexes or four-way DNA junctions, which share the architectural feature of presenting two double strands emerging from a central bend [36]. Indeed, the characterization of an XPA mutant, where neighboring lysines in a positively charged DNA-binding surface were replaced by glutamates, indicates that the assembly of active incision complexes is dependent on the close association of XPA protein with a narrow bending angle in the DNA substrate [38].

How does an affinity for kinked DNA contribute to damage recognition? Base stacking is the predominant energetic force leading to the intrinsic rigidity of DNA but the loss of base stacking, resulting in the formation of dynamic hinge points, is a common consequence of DNA lesions [16]. This property of damaged DNA implies that XPA protein may exploit its affinity for sharply bent DNA backbones to carry out a damage verification function by

another indirect mechanism, i.e., by probing the susceptibility of the DNA substrate to be kinked during assembly of the lesion verification complex [38].

Conclusions and perspectives

One of the most formidable challenges in molecular recognition is that faced by the initiators of the NER reaction as they locate damaged bases among a large excess of undamaged residues. This challenge is further complicated by the amazing diversity of target lesions, and it is evident that substrate recognition in the NER pathway must display high levels of versatility. An intriguing new aspect emerging from the molecular analysis of XPC and other mammalian NER subunits is that a subset of early recognition factors operate in an entirely indirect manner by detecting damage-induced alterations in DNA dynamics, including increased strand oscillations and the susceptibility of the DNA backbone to undergo site-specific kinks. The notion that generic DNA damage recognition factors may operate in a strictly indirect manner, thus avoiding close contacts with injured bases, is supported by recent studies on the prokaryotic NER system [19,39,40].

How does this novel concept of damage recognition guided by abnormal DNA dynamics, such as for example strand oscillations, change the direction of research in the NER field? Clearly, future experiments should not be limited to the analysis of static molecular complexes formed by DNA damage recognition factors and their nucleic acid substrates. Instead, new experimental approaches will be aimed at the measurement of dynamic changes induced by biologically significant DNA lesions (see Box 2). Quantitative relationships will be established between the type and degree of such dynamic changes and the efficiency of lesion recognition as well as the rate of NER complex assembly. Advanced real-time methods will be employed in reconstituted systems and in intact cells to monitor the recruitment of NER factors and the specific reciprocal interactions arising between multiple

subunits in response to alterations of DNA dynamics. Another important issue will be to examine how reversible posttranslational modifications of NER subunits may facilitate the recognition of transient dynamic changes of the DNA substrate.

Acknowledgements

H.N. is supported by the Swiss National Science Foundation (grant 3100A0-113694) and by Oncosuisse (grant KLS-01827-02-2006). The work of K.B.B. is partly supported by the Center for Nonlinear Studies at the Los Alamos National Laboratory.

Box 1. The human NER reaction and its core subunits

Many human NER proteins are encoded by genes that, when mutated, give rise to xeroderma pigmentosum (XP), an inherited disorder characterized by extreme photosensitivity and a 2000-fold increased incidence of sunlight-induced skin cancer [1,41]. The XP syndrome further involves a higher risk of contracting internal tumors and, in some cases, neurological complications, presumably reflecting the essential role of NER in the removal of endogenous DNA damage. The products of the *XPA-XPG* genes provide the core subunits of an excision complex that comprises approximately 20 polypeptide subunits.

The different NER steps, i.e., recognition of DNA distortions, verification of base lesions, dual DNA incision and repair synthesis, are illustrated in the accompanying scheme (Figure I), where the substrate carries a benzo[a]pyrene diol epoxide adduct. The NER system excises such carcinogen-DNA adducts, as well as UV photoproducts or other lesions, in the form of oligonucleotide segments that have a length of 24-32 residues [42]. This reaction is generally thought to proceed by a stepwise mechanism initiated by a heterotrimeric factor consisting of XPC, RAD23B and CETN2 (centrin-2) [13,14,43]. XPC protein possesses DNA-binding activity, whereas the RAD23B and CETN2 partners exert accessory functions by stabilizing the complex and stimulating its action in DNA repair. UV-damaged DNA-binding (UV-DDB) protein may accelerate the recognition of UV photoproducts [8]. After the initial association with damaged sites, XPC protein mediates the recruitment of transcription factor IIH (TFIIH), followed by replication protein A (RPA), XPA, XPG and, finally, XPF-ERCC1, which is a heterodimer composed of XPF and excision repair cross complementing-1 [44,45].

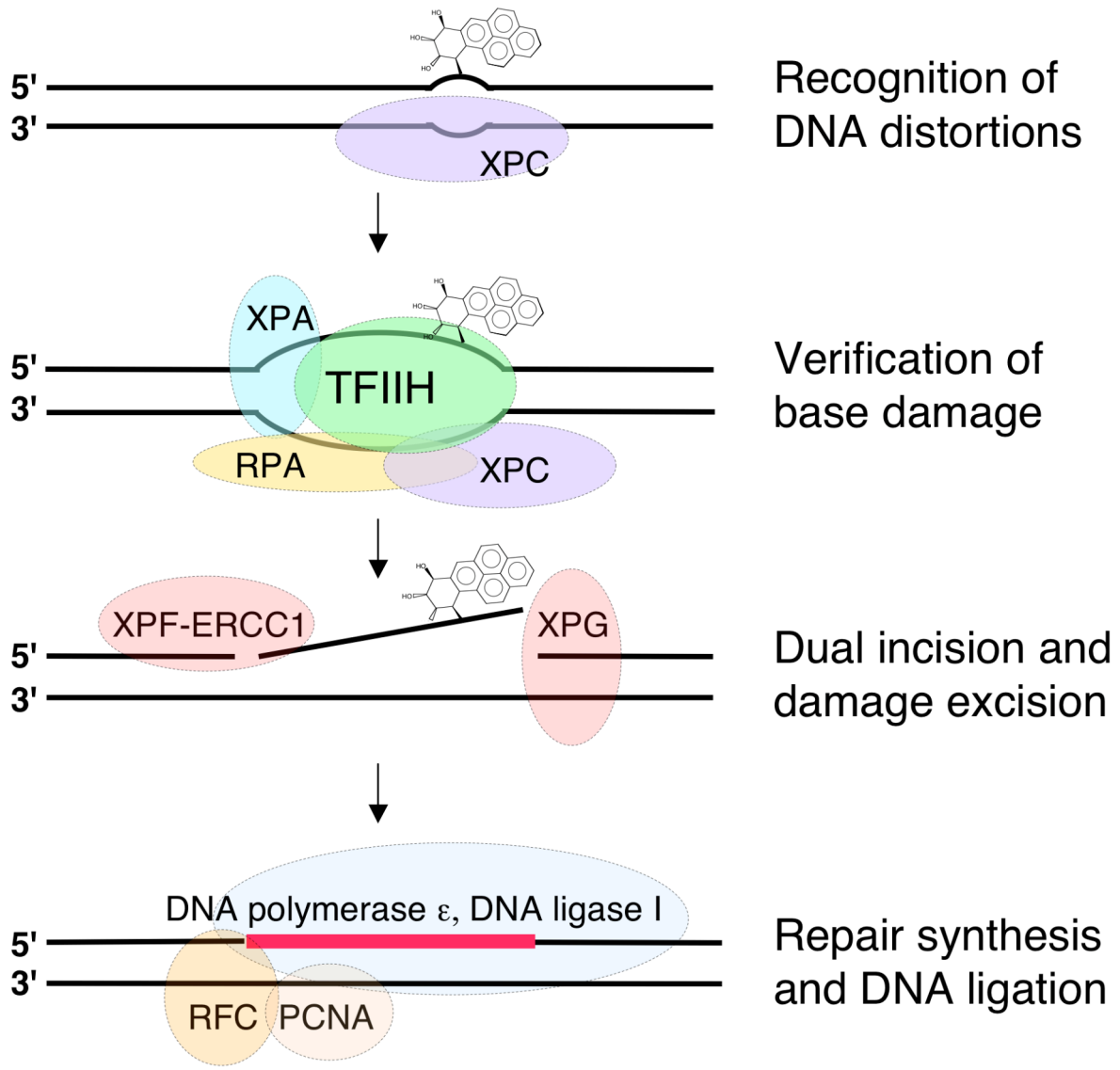


Figure Box I Human NER pathway and its core subunits. The damaged strand of the DNA substrate carries a polycyclic aromatic hydrocarbon adduct. See text for abbreviations.

Four of these early factors (XPC, TFIIH, RPA and XPA) generate a unique recognition intermediate characterized by transient unwinding of the duplex substrate around the lesion. This open intermediate contains "Y-shaped" double- to single-stranded DNA transitions, which provide a substrate for the junction-specific DNA endonucleases XPF-ERCC1 and

XPG [46,47]. By cleaving the damaged strand at the "Y-shaped" junctions, these two endonucleases act as "scissors" to cut out DNA damage from the double helix. XPF-ERCC1 makes the 5' and XPG the 3' incision. Helical integrity is reconstituted by a downstream DNA synthesis complex consisting of an error-free DNA polymerase (primarily DNA polymerase ε), in conjunction with replication factor C (RFC) and proliferating cell nuclear antigen (PCNA). Subsequently, the newly synthesized DNA repair patches are joined to the preexisting strands by the action of DNA ligase I [44,48].

Box 2. Outstanding questions

How does the degree of DNA oscillations relate to repair efficiency?

Nuclear magnetic resonance (NMR) and optical correlation spectroscopy (OCS) will be used to measure changes of helical dynamics resulting from the formation of DNA damage. These methods involve a magnetic label in the case of NMR, or fluorescent labels for OCS. The NMR analyses are limited to small nucleic acid molecules in the range of 30 base pairs, but OCS can also be applied to the study of larger DNA duplexes. The frequency and amplitude of strand oscillations can be modulated by isotope substitutions in the DNA base pairs. Thereafter, the degree of strand oscillations or other DNA fluctuations will be compared to the rates of NER assembly and efficiency of damage excision.

How does UV-DDB mediate the recognition of cyclobutane pyrimidine dimers?

An accessory factor (UV-DDB) is thought to facilitate the recognition of cyclobutane pyrimidine dimers, the most frequent photoproduct induced by exposure to sunlight [8]. Intriguingly, UV-DDB seems to be necessary for the recruitment of NER factors to cyclobutane pyrimidine dimers in intact cells but not in reconstituted *in vitro* systems [7]. Thus, although UV-DDB may represent the first factor that recognizes pyrimidine dimers in human cells, its mode of interaction with damaged sites, as well as the physical handover of DNA lesions from UV-DDB to the next NER factors, needs to be analyzed in the chromosomal context.

What is the role of protein ubiquitylation in the NER pathway?

UV-DDB is not only a damage recognition subunit but also a molecular adaptor that connects the Cul4A-Roc1 ubiquitin ligase complex to a wide repertoire of protein targets. Known

substrates of this ubiquitylation machinery include, in addition to UV-DDB itself, various histones and XPC protein [49]. The ubiquitylation reaction appears to stimulate excision activity but it is not clear why the presence of small polypeptide modifiers should be important for either damage recognition or the assembly of excision complexes.

How do the other NER subunits contribute to DNA damage recognition?

The NER pathway is likely to function as a recognition cascade, in which a sequence of multiple steps of mediocre selectivity ultimately yields an excellent specificity for DNA lesions [9,32]. However, it is not yet clear how the downstream factors TFIIH, RPA, XPG or XPF-ERCC1 contribute to this selectivity cascade.

Damage recognition in the transcription-coupled pathway

XPC and UV-DDB are not always needed for NER activity, as these factors are dispensable for the transcription-coupled subpathway [1,50]. How the transcription complex mediates DNA damage recognition and recruits the repair machinery are other unsolved problems.

References

1 Friedberg, E.C. *et al*. (2006) *DNA Repair and Mutagenesis*, ASM Press

2 Reardon, J.T. *et al*. (1997) In vitro repair of oxidative DNA damage by human nucleotide excision repair system: possible explanation for neurodegeneration in xeroderma pigmentosum patients. *Proc. Natl. Acad. Sci. USA* 94, 9463-9468

3 Kuraoka, I. *et al*. (2000) Removal of oxygen free-radical-induced 5',8-purine cyclodeoxynucleosides from DNA by the nucleotide excision-repair pathway in human cells. *Proc. Natl. Acad. Sci. USA* 97, 3832-3837

4 Reardon, J.T. and Sancar, A. (2006) Repair of DNA-polypeptide crosslinks by human excision nuclease. *Proc. Natl. Acad. Sci. USA* 103, 4056-4061

5 Huang, J.C. *et al*. (1994) Substrate spectrum of human excinuclease: repair of abasic sites, methylated bases, mismatches and bulky adducts. *Proc. Natl. Acad. Sci. USA* 91, 12213-12217

6 Mu, D. *et al*. (1997) Characterization of reaction intermediates of human excision repair nuclease. *J. Biol. Chem.* 272, 28971-28979

7 Reardon, J.T. and Sancar, A. (2003) Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by the human excision nuclease. *Genes Dev.* 17, 2539-2551

8 Tang, J.Y. *et al*. (2000) Xeroderma pigmentosum p48 gene enhances global genomic repair and suppresses UV-induced mutagenesis. *Mol. Cell.* 5, 737-744

9 Volker, M. *et al*. (2001) Sequential assembly of the nucleotide excision repair factors in vivo. *Mol. Cell* 8, 213-224

10 Evans, E. *et al*. (1997) Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO J.* 16, 6559-6573

11 Riedl, T. *et al*. (2003) The comings and goings of nucleotide excision repair factors on damaged DNA. *EMBO J.* 22, 5293-5303

12 Politi, A. *et al*. (2005) Mathematical modeling of nucleotide excision repair reveals efficiency of sequential assembly strategies. *Mol. Cell* 19, 679-690

13 Sugasawa, K. *et al*. (1998) Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Mol. Cell* 2, 223-232

14 Hanawalt, P.C. and Haynes, R.H. (1965) Repair replication of DNA in bacteria: irrelevance of chemical nature of base defect. *Biochem. Biophys. Res. Commun.* 19, 462-467

15 Gunz, D. *et al*. (1996) Recognition of DNA adducts by human nucleotide excision repair. Evidence for a thermodynamic probing mechanism. *J. Biol. Chem.* 271, 25089-25098

16 Isaacs, R.J. and Spielmann, H.P. (2004) A model for initial DNA lesion recognition by NER and MMR based on local conformational flexibility. *DNA Rep.* 3, 455-464

17 Marx, A. *et al*. (1997) 4'-acetylated thymidines: a new class of DNA chain terminators and photocleavable DNA building blocks. *J. Am. Chem. Soc.* 119, 1131-1132

18 Buterin, T. *et al*. (2005) DNA quality control by conformational readout on the undamaged strand of the double helix. *Chem. Biol.* 12, 913-922

19 Malta, E. *et al.* (2006) Base flipping in nucleotide excision repair. *J. Biol. Chem.* 281, 2184-2194

20 Alexandrov, B.S. *et al*. (2006) Bubble statistics and dynamics in double-stranded DNA. *Phys. Rev. E* 74, 05090

21 Blagoev, K.B. *et al*. (2006) Ultra-violet light induced changes in DNA dynamics may enhance TT-dimer recognition. *DNA Rep.* 5, 863-867

22 Cosman, M. *et al*. (1996) Solution conformation of the (-)-*cis*-*anti*-benzo[*a*]pyrenyl-dG adduct opposite dC in a DNA duplex: intercalation of the covalently attached BP ring into

the helix with base displacement of the modified deoxyguanosine into the major groove. *Biochemistry* 35, 9850-9863

23 Bertrand-Burggraf, E. *et al*. (1994) Endonuclease VII of phage T4 nicks *N*-2-acetylaminofluorene-induced DNA structures in vitro. *Mutat. Res*. 314, 287-295

24 Batty, D. *et al*. (2000) Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged sites. *J. Mol. Biol.* 300, 275-290

25 Sugasawa, K. *et al*. (2001) A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes Dev.* 15, 507-521

26 Maillard, O. *et al*. (2007) An aromatic sensor with aversion to damaged strands confers versatility to DNA repair. *PLoS Biol.* 5, e79

27 Bochkarev, A. *et al*. (1997) Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature* 385, 176-181

28 Trego, K.S. and Turchi, J.J. (2006) Pre-steady-state binding of damaged DNA by XPC-hHR23B reveals a kinetic mechanism for damage discrimination. *Biochemistry* 45, 1961-1969

29 Shimizu, Y. *et al*. (2003) Xeroderma pigmentosum group C protein interacts physically and functionally with thymine DNA glycosylase. *EMBO J.* 22, 164-173

30 D'Errico, M. *et al*. (2006) New functions of XPC in the protection of human skin cells from oxidative damage. *EMBO J.* 25, 4305-4315

31 Despras, E. *et al*. (2007) Long-term XPC silencing reduces DNA double-strand break repair. *Cancer Res.* 67, 2526-2534

32 Lin, J.J. and Sancar, A. (1992) (A)BC excinuclease: the Escherichia coli nucleotide excision repair enzyme. *Mol. Microbiol.* 6, 2219-2224

33 Yokoi, M. *et al*. (2000) The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIH to damaged DNA. *J. Biol. Chem.* 275, 9870-9875

34 Naegeli, H. *et al*. (1992) The DNA helicase and adenosine triphosphatase activities of yeast Rad3 protein are inhibited by DNA damage. *J. Biol. Chem.* 267, 392-398

35 Fan, L. *et al*. (2006) Conserved XPB core structure and motifs for DNA unwinding: implications for pathway selection of transcription or excision repair. *Mol. Cell* 22, 27-37

36 Missura, M. *et al*. (2001) Double-check probing of DNA bending and unwinding by XPA-RPA: an architectural function in DNA repair. *EMBO J.* 20, 3554-3564

37 Li, L. *et al*. (1995) An interaction between the DNA repair factor XPA and replication protein A appears essential for nucleotide excision repair. *Mol. Cell. Biol.* 15, 5396-5402

38 Camenisch, U. *et al*. (2006) Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair. *Nat. Struct. Mol. Biol.* 13, 278-284

39 Truglio, J.J. *et al.* (2006) Structural basis for DNA recognition and processing by UvrB. *Nat. Struct. Mol. Biol.* 13, 360-364

40 Waters, *T.R. et al.* (2006) Damage detection by the UvrABC pathway: crystal structure of UvrB bound to fluorescein-adducted DNA. *FEBS Lett.* 580, 6423-6427

41 Kraemer, K.H. *et al*. (1984) DNA repair protects against cutaneous and internal neoplasia: evidence from xeroderma pigmentosum. *Carcinogenesis* 5, 511–514

42 Huang, J.C. *et al*. (1992) Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer. *Proc. Natl. Acad. Sci. USA* 89, 3664-3668

43 Nishi, R. *et al*. (2005) Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein. *Mol. Cell. Biol.* **25**: 5664-5674

44 Aboussekhra, A. *et al*. (1995) Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell* 80, 859-868

45 Mu, D. *et al*. (1995) Reconstitution of human DNA repair excision nuclease in a highly defined system. *J. Biol. Chem.* 270, 2415-2418

46 Evans, E. *et al*. (1997) Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein. *EMBO J.* 16, 625-638

47 Bessho, T. *et al*. (1997) Reconstitution of human excision nuclease with recombinant XPF-ERCC1 complex. *J. Biol. Chem.* 272, 3833-3837

48 Shivji, M.K. *et al*. (1995) Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA. *Biochemistry* 34, 5011-5017

49 Sugasawa, K. *et al*. (2005) UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex. *Cell* 121, 387-400

50 Andressoo, J.O. and Hoeijmakers, J.H. (2005) Transcription-coupled repair and premature ageing. *Mutat. Res.* 577, 179-194

Summary

DNA, the carrier of our genetic information, is constantly challenged by different genotoxic agents such as, for example, UV-light, carcinogenic compounds or endogenous reactive metabolites. The resulting DNA lesions can, in turn, perturb vital cellular processes such as transcription, replication and cell cycle progression. Persistent DNA damage can induce mutations and ultimately lead to cancer or to cell death. In order to maintain the integrity of the genetic material, cells have evolved an intricate network of DNA repair mechanisms, which are shortly summarized in **Chapter 3**.

One of the primary DNA repair pathways is nucleotide excision repair. It is responsible for the removal of a wide spectrum of bulky DNA lesions that have in common their ability to distort the helical structure of double-stranded DNA. These lesions include photoproducts induced by UV light and DNA adducts generated by carcinogenic chemicals. The nucleotide excision repair pathway requires the coordinated activity of approximately 30 different proteins. The mechanism involves a multistep "cut and patch" reaction in which a short segment of DNA containing the lesion is excised and replaced by DNA repair synthesis using the intact complementary strand as the template. Two distinct nucleotide excision repair subpathways can be distinguished; i.e. global genome repair (GGR), operating over the entire genome, and transcription-coupled repair (TCR), which repairs only the transcribed strand of active genes. The principal difference between these pathways resides in the initial detection of DNA damage. During TCR, elongation of the RNA polymerase II complex is blocked by abnormal residues, thereby inducing the assembly of repair complexes. In contrast, the GGR machinery is dependent on the initial recognition of damaged sites by the XPC-protein. The human nucleotide excision repair pathway and its subunits are extensively discussed in **Chapter 3 & 4**.

Xeroderma pigmentosum group C (XPC) protein is the key damage recognition factor that initiates the GGR process. However, the molecular mechanism by which this versatile sensor of DNA damage detects a wide range of bulky base lesions was not understood. For that reason, during the course of my thesis I investigated the substrate recognition function of human XPC protein in more detail. The starting point and the goals of my thesis project are specified in **Chapter 1**.

In a first step, I examined the substrate recognition function of XPC by subjecting evolutionary conserved amino acids to site-directed mutagenesis. After screening for

repair activity in a host cell reactivation assay, the repair-defective mutants were analyzed for expression level, nuclear localization and their DNA binding capacity. This strategy revealed that two particular aromatic amino acids (tryptophan at position 690 and phenylalanine at position 733) are critically involved in the interaction of XPC protein with its DNA substrate. In parallel, biochemical experiments demonstrated that XPC protein displays a strong preference for binding to undamaged single-stranded DNA. The aromatic residues at codon 690 and 733 are absolutely required for this interaction with single-stranded DNA. Interestingly, the dual deployment of aromatic side chains for the interaction with nucleic acids is the distinctive feature of oligonucleotide/oligosaccharide-binding folds found in single-stranded DNA binding proteins such as for example replication protein A. The detailed results of my thesis work are outlined in **Chapter 2**.

On the basis of these results, we concluded that human XPC protein exhibits an entirely inverted mechanism of substrate recognition whereby this subunit avoids direct contacts with the damaged bases themselves. Instead, XPC protein probes the local susceptibility of intact nucleotides on the opposite undamaged side of the double helix to adopt a single-stranded configuration. This indirect mechanism of substrate recognition accounts for the exceptional substrate versatility of this initial damage recognition sensor.

Zusammenfassung

Die in der DNA – unserem Erbgut – gespeicherte Information ist stetig unterschiedlichen genotoxischen Einflüssen - wie zum Beispiel UV-Licht, krebserregenden Chemikalien oder reaktiven Stoffwechselprodukten - ausgesetzt, welche die DNA schädigen. Diese Schäden führen kurzfristig zu Störungen der lebenswichtigen zellulären Prozesse wie zum Beispiel der Transkription, der Replikation oder des Zellzyklus. Bleibende DNA Schäden können zu Mutationen führen, welche ihrerseits wiederum Erbkrankheiten, Krebs oder den Zelltod verursachen können. Um genau dies zu verhindern und die Integrität der DNA zu schützen, haben unsere Zellen im Laufe der Evolution ein komplexes Netzwerk verschiedener DNA-Reparaturmechanismen entwickelt. Diese sind in **Kapitel 3** kurz zusammengefasst.

Eines der wichtigen DNA-Reparatursysteme ist die so genannte „Nukleotid-Exzisionsreparatur“. Sie ist verantwortlich für die Behebung eines grossen Spektrums von DNA-Schäden, deren einzige Gemeinsamkeit die Fähigkeit ist, die helikale Struktur der DNA zu stören. Beispiele solcher Schäden sind durch UV-Licht induzierte Photoprodukte oder durch Chemikalien verursachte „unförmige“ DNA-Addukte. Die Durchführung der Nukleotid-Exzisionsreparatur benötigt die koordinierte Aktivität von etwa 30 verschiedenen Proteinen. Der Mechanismus umfasst ein mehrstufiges „Ausschneiden und Einfügen“, wobei ein kurzes Segment des beschädigten DNA-Stranges herausgeschnitten und durch eine neu synthetisierte Kopie des intakten komplementären Stranges ersetzt wird. Die Nukleotid-Exzisionsreparatur wird in zwei verschiedene Untersysteme eingeteilt. Zum einen in die „globale genomische Reparatur“ (GGR), welche Schäden im ganzen Genom behebt, und zum anderen in die so genannte „trankriptionsgekoppelte Reparatur“ (transcription-coupled repair; TCR), welche spezifisch Schäden an der aktuell zu transkribierenden DNA behebt. Der Hauptunterschied der beiden Systeme liegt in der Initiierung des Prozesses, das heisst im Mechanismus der initialen Schadenserkennung. Im Falle der TCR wird die Reparatur von einem durch eine Läsion blockierten RNA-Polymerase-II-Komplex initiiert. Im Unterschied dazu ist die GGR vollständig von der Schadenserkennung durch das XPC-Protein abhängig. Das System der Nukleotid-Exzisionsreparatur und die daran beteiligten Proteine werden ausführlich in **Kapitel 3 & 4** besprochen.

Im GGR-Prozess ist das XPC-Protein der Schlüssel-Faktor, der für die Schadenserkennung und die Initiierung der DNA-Reparatur verantwortlich ist. Der molekulare Mechanismus, mit welchem dieser vielseitige Sensor von DNA-Schäden eine so reichhaltige Palette von unterschiedlichen Schäden erkennen kann, war jedoch nicht bekannt. Deshalb habe ich in meiner Arbeit die Schadenserkennungs-Funktion des XPC-Proteins eingehend untersucht. In **Kapitel 1** sind die Ausgangslage und die Ziele meiner Arbeit ausführlich beschrieben.

In einem ersten Schritt untersuchte ich den Einfluss von evolutionär konservierten Aminosäuren, welche ich mit zielgerichteter Mutagenese veränderte, in Bezug auf die Schadenserkennung durch das humane XPC-Protein. Mit Hilfe eines so genannten Wirtszell-Reaktivierungsassays wurden die Mutanten auf ihre Reparaturaktivität hin untersucht. Anschliessend untersuchte ich den Expressionslevel, die nukleäre Lokalisation und die DNA-Bindung der Mutanten, die im funktionellen Assay einen bedeutenden Reparaturdefekt aufwiesen. Diese Strategie offenbarte zwei kritische aromatische Aminosäuren (Tryptophan an der Position 690 und Phenylalanin an der Position 733), welche an der Interaktion des XPC-Proteins mit seinem DNA-Substrat beteiligt sind. Daneben zeigten biochemische Bindungassays, dass das XPC-Protein bevorzugt an ungeschädigte Einzelstrang-DNA bindet und dass die beiden Aromaten (Tryptophan690 und Phenylalanin733) für diese Wechselwirkung mit einzelsträngigen DNA notwendig sind. Exakt dieses Aminosäuremuster mit zwei herausragenden aromatischen Seitenketten, findet sich auch im so genannten Oligonukleotid/Oligosaccharid-Bindungsmotiv wieder, welches Bestandteil von DNA-Einzelstrangbindungsproteinen wie zum Beispiel dem Replikationsprotein A ist. Tatsächlich weist der betreffende Abschnitt des humanen XPC-Proteins eine auffällige Sequenz-Homologie mit diesem Oligonukleotid/Oligosaccharid-Bindungsmotiv auf. Die experimentellen Ergebnisse meiner Dissertation finden sich in **Kapitel 2**. Aufgrund meiner Befunde drängte sich die absolut neuartige Hypothese auf, dass das XPC-Protein gar nicht direkt an den Schaden bindet, sondern dass es die einzelsträngige Konformation der DNA-Doppelhelix im ungeschädigten, komplementären Strang erkennt. Durch diesen indirekten aber äusserst vielseitigen Mechanismus ist erklärbar, dass ein einzelnes Sensorprotein die Reparatur einer sehr breiten Auswahl von DNA-Schäden einleiten kann.

Curriculum Vitae

Particulars

Surname	MAILLARD
First Names	Olivier Nicolas
Date of Birth	12.03.1975
Nationality	Cressier FR, Switzerland

Education

High school	Deutsches Gymnasium Biel Graduation 1995, Typus C
Studies	University of Bern, Veterinary medicine 4 Semesters, 1995 to 1997
Graduate studies	Biology with major in microbiology 8 Semesters, 1997 to 2001
Diploma thesis	„Function of Mag1 of *Schizosaccharomyces pombe* in DNA mismatch repair“ Institute of Cell-Biology, University of Bern Dr. Oliver Fleck, December 2001
Stage	Follow up of Diploma thesis Institute of Cell-Biology, University of Bern Dr. Oliver Fleck, till September 2002
PhD thesis	University of Zürich, Vetsuisse Institute of Veterinary Pharmacology and Toxicology Prof. Dr. Hanspeter Naegeli Since October 2002

Meetings

Poster	Meeting on DNA repair: from Molecular mechanism to Human Disease Noordwijkerhout, The Netherlands, April 2-7, 2006
Talk	2nd Swiss Meeting on genome Stability DNA Dynamics and Epigenetics Zehntenhaus – Uetendorf, October 12-14, 2005
	15. Vetpharm Symposium Vienna, Austria, September 29-30, 2005

List of Publications

Maillard O, Solyom S, Naegeli H (2007) An Aromatic Sensor with Aversion to Damaged Strands Confers Versatility to DNA Repair. *PLoS Biology* Vol. 5, No. 4, e79 doi:10.1371/journal.pbio.0050079

Olivier Maillard, Ulrike Camenisch, Flurina C. Clement, Krastan B. Blagoev and Hanspeter Naegeli (2007) DNA repair triggered by sensors of helical dynamics. *TiBS*, in Press Accepted, 29. August 2007; Ref.: Ms. No. TIBS-D-07-00027R2

Olivier Maillard, Ulrike Camenisch, Krastan B. Blagoev and Hanspeter Naegeli (2007) Versatile protection from mutagenic DNA lesions conferred by bipartite recognition in nucleotide excision repair (submitted to Mutation Resaerch)

Acknowledgements

I would like to thank cordially:

Prof. Dr. Hanspeter Naegeli for the possibility to realise my PhD in his group, for the excellent supervision of my work and his great support in professional and personal regards
Prof. Dr. Josef Jiricny for taking over the chair of the thesis committee.
Prof. Dr. Michael Hottiger for joining the thesis committee.
Prof. Dr. Ulrich Hübscher and Prof. Dr. Benjamin Schuler for critical reading of my thesis and writing the experts opinion reports.
Prof. Dr. Felix Althaus for helpful and suggestive discussions.

All my colleagues from the Institute of Veterinary Pharmacology and Toxicology:
I was able to spend an absolutely great time in a really nice environment. We had a lot of fun during work and our „social activities" (partys, lakeside partys, movie evenings, diners, sports,...).

In particular:
Kiki Camenisch, Flurina Clement & Odile Cohausz for having a great time not just in but also beside the lab with fantastic diners, great and long party's and cinema evenings.
Mirela Vitanescu, Müriel Träxler & Maja Bollhalder for their support in daily lab work
Daniel Demuth & Cedric Müntener for their support with computer matters
Susanne Bachmann for her support with administrative affairs

Last but not least a very spezial thank goes to

My **parents, Roland & Susanne Maillard,** who always believed in me and supported me wherever and whenever they could. Thanks for everything!
My **sister, her husband and their daughter, Annick & Stefan & Shoana-Alyssa Leu-Maillard** for their familial support.

The Swiss National Science Foundation financially supported this work